

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 18:13:00 ; Search time 162 Seconds
(without alignments)
35.811 Million cell updates/sec

Title: US-09-623-038-6

Perfect score: 87

Sequence: 1 RSYQHDLRAYGFWR 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: Genesep1980s:*

2: Genesep1990s:*

3: Genesep2000s:*

4: Genesep2001s:*

5: Genesep2002s:*

6: Genesep2003as:*

7: Genesep2003bs:*

8: Genesep2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------|-------------|
| 1 | 87 | 100.0 | 15 | AAV30352 | Epitope d |
| 2 | 87 | 100.0 | 15 | AAE19240 | Streptococ |
| 3 | 87 | 100.0 | 66 | AAE19245 | Bi-peptid |
| 4 | 82 | 94.3 | 15 | AAE19244 | Streptococ |
| 5 | 82 | 94.3 | 47 | AAE19247 | Tri-pepti |
| 6 | 47 | 54.0 | 554 | ABG19785 | Novel hum |
| 7 | 47 | 54.0 | 554 | ABG19731 | Novel hum |
| 8 | 47 | 54.0 | 1007 | ABG03083 | Novel hum |
| 9 | 45 | 51.7 | 741 | ABP79868 | N. gonorr |
| 10 | 45 | 51.7 | 903 | ADF03901 | Bacterial |
| 11 | 44 | 50.6 | 118 | ABP02252 | Human ORF |
| 12 | 44 | 50.6 | 267 | ABG26189 | Novel hum |
| 13 | 44 | 50.6 | 1096 | AAE25282 | Human nuc |
| 14 | 44 | 50.6 | 1114 | AAE08517 | Amino aci |
| 15 | 44 | 50.6 | 1262 | AAE08463 | Human pro |
| 16 | 44 | 50.6 | 1289 | AAE079447 | Human pro |
| 17 | 44 | 50.6 | 1356 | ADC06830 | Human pro |
| 18 | 44 | 50.6 | 1593 | ADC06829 | Human pro |
| 19 | 44 | 50.6 | 1891 | ABG04331 | Novel hum |
| 20 | 43 | 49.4 | 29 | ABP40987 | Human ret |
| 21 | 43 | 49.4 | 126 | ABP80192 | N. gonorr |
| 22 | 43 | 49.4 | 360 | ABP40989 | Human ret |
| 23 | 43 | 49.4 | 420 | ABP40990 | Human ret |
| 24 | 43 | 49.4 | 433 | ABP78342 | N. gonorr |
| 25 | 43 | 49.4 | 459 | AAV75225 | Neisseria |

| | | | | | | |
|----|----|------|-----|---|----------|-----------|
| 26 | 43 | 49.4 | 459 | 3 | AAV75223 | Neisseria |
| 27 | 43 | 49.4 | 465 | 6 | ABU63641 | Bsmal res |
| 28 | 43 | 49.4 | 466 | 6 | ABU32026 | Protein e |
| 29 | 43 | 49.4 | 484 | 7 | ABO63378 | Klebsiell |
| 30 | 43 | 49.4 | 530 | 6 | ABR43482 | Bacillus |
| 31 | 42 | 48.3 | 280 | 6 | ABM68845 | Phototrab |
| 32 | 42 | 48.3 | 652 | 4 | ABM67974 | Drosophil |
| 33 | 42 | 48.3 | 951 | 2 | AAV34536 | Porphorym |
| 34 | 42 | 48.3 | 953 | 2 | AAV34403 | Porphorym |
| 35 | 41 | 47.1 | 111 | 4 | ABG10728 | Novel hum |
| 36 | 41 | 47.1 | 111 | 4 | ABG25967 | Novel hum |
| 37 | 41 | 47.1 | 144 | 7 | ADF05142 | Bacterial |
| 38 | 41 | 47.1 | 227 | 7 | ABO76461 | Pseudomon |
| 39 | 41 | 47.1 | 497 | 7 | ADI27244 | B. napus |
| 40 | 41 | 47.1 | 544 | 8 | ABG98169 | Bsal rest |
| 41 | 41 | 47.1 | 588 | 4 | AAU27909 | Human con |
| 42 | 41 | 47.1 | 651 | 4 | ABG08359 | Novel hum |
| 43 | 41 | 47.1 | 692 | 8 | ADJ49464 | Oil-assoc |
| 44 | 41 | 47.1 | 817 | 8 | ADJ49477 | Oil-assoc |
| 45 | 40 | 46.0 | 87 | 4 | ABG08529 | Novel hum |

ALIGNMENTS

RESULT 1

AAV30352

ID AAV30352 standard; peptide; 15 AA.

XX AAV30352;

XX

XX 09-NOV-1999 (first entry)

XX Epitope derived from pneumococcal surface adhesion A protein.

XX Pneumococcal surface adhesion A protein; PsaA; monoclonal antibody;

XX vaccine; Streptococcus pneumoniae infection.

XX Streptococcus pneumoniae.

XX WO9945121-A1.

XX 10-SEP-1999.

XX 26-FEB-1999; 99WO-US004326.

XX 02-MAR-1998; 98US-0076565P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Carlone GM, Ades EW, Sampson JS, Tharpe JA, Zeiler JL;

XX Westerink MAJ;

XX WPI; 1999-540849/45.

XX New peptides corresponding to Streptococcus pneumoniae PsaA, used for

XX treating or preventing Streptococcus pneumoniae infection in a subject.

XX Claim 6; Page 43; 58pp; English.

XX AAV30351-54 represent immunogenic peptides which are derived from a

XX pneumococcal surface adhesion A protein (PsaA). The specification

XX describes monoclonal antibodies which bind epitopes of the PsaA protein

XX (e.g present sequence). The peptides can be used in vaccines to prevent

XX Streptococcus pneumoniae infections. The antibodies of the invention can

XX also be used to detect S. pneumoniae in a sample or individual

XX SQ Sequence 15 AA;

XX Query Match 100.0%; Score 87; DB 2; Length 15;

XX Best Local Similarity 100.0%; Pred. No. 1.7e-07;

XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 RSYQHDLRAYGFWR 15
Db 1 RSYQHDLRAYGFWR 15

RESULT 2
AAE19240
ID AAE19240 standard; peptide; 15 AA.
XX
XX AAE19240;
XX
XX 21-MAY-2002 (first entry)
XX
XX Streptococcus pneumoniae PsaA immunogenic peptide #2.
XX Multiple antigenic peptide; MAP; immunogenic; immunity; infection;
XX pneumococcal surface adhesin protein A; PsaA; antibacterial.
XX
XX Streptococcus pneumoniae.
XX
XX WO200204497-A2.
XX
XX 17-JAN-2002.
XX
XX 10-JUL-2001; 2001WO-US021626.
XX
XX 10-JUL-2000; 2000US-00613092.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Ades EW, Johnson SE, Jue DL, Sampson JS, Carlone GM;
XX WPI; 2002-195762/25.
XX
XX New multiple antigenic peptide for immunizing against streptococcal
XX infections, binds to monoclonal antibody obtained in response to
XX immunizing an animal with pneumococcal surface adhesion protein A or its
XX fragment.
XX
XX Claim 2; Page 56; 86pp; English.
XX
XX The invention relates to multiple antigenic peptides (MAP) immunogenic
XX against Streptococcus pneumoniae. MAP binds to monoclonal antibody
XX obtained in response to immunising an animal with pneumococcal surface
XX adhesion protein A (PsaA) or its fragment. MAP is useful for conferring
XX protective immunity against S. pneumoniae infection in a subject. The
XX present sequence is a bi-peptide heterogenous MAP used to provide
XX protection against Streptococcus pneumoniae challenge
XX
XX Sequence 15 AA;
XX
XX Query Match 100.0%; Score 87; DB 5; Length 15;
XX Best Local Similarity 100.0%; Pred. No. 1.7e-07;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX OY 1 RSYQHDLRAYGFWR 15
XX Db 1 RSYQHDLRAYGFWR 15
XX
XX RESULT 3
XX AAE19245
XX ID AAE19245 standard; protein; 66 AA.
XX
XX AC AAE19245;
XX
XX 21-MAY-2002 (first entry)
XX
XX Bi-peptide heterogenous multiple antigenic peptide (MAP) #1.
XX Multiple antigenic peptide; MAP; immunogenic; immunity; infection;
XX pneumococcal surface adhesin protein A; PsaA; antibacterial.
XX
XX Unidentified.
XX
XX WO200204497-A2.

XX Key Location/Qualifiers
XX Peptide 1..15
XX /note= "PsaA immunogenic peptide #1"
XX Peptide 16..30
XX /note= "PsaA immunogenic peptide #2"
XX Peptide 31..45
XX /note= "PsaA immunogenic peptide #1"
XX Peptide 46..60
XX /note= "PsaA immunogenic peptide #2"
XX Modified-site 62
XX /label= Nle
XX Modified-site 64
XX /label= Nle
XX Modified-site 66
XX /label= Nle
XX WO200204497-A2.
XX 17-JAN-2002.
XX 10-JUL-2001; 2001WO-US021626.
XX 10-JUL-2000; 2000US-00613092.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX Ades EW, Johnson SE, Jue DL, Sampson JS, Carlone GM;
XX WPI; 2002-195762/25.
XX
XX New multiple antigenic peptide for immunizing against streptococcal
XX infections, binds to monoclonal antibody obtained in response to
XX immunizing an animal with pneumococcal surface adhesion protein A or its
XX fragment.
XX
XX Example 1b; Fig 1A; 86pp; English.
XX
XX The invention relates to multiple antigenic peptides (MAP) immunogenic
XX against Streptococcus pneumoniae. MAP binds to monoclonal antibody
XX obtained in response to immunising an animal with pneumococcal surface
XX adhesion protein A (PsaA) or its fragment. MAP is useful for conferring
XX protective immunity against S. pneumoniae infection in a subject. The
XX present sequence is a bi-peptide heterogenous MAP used to provide
XX protection against Streptococcus pneumoniae challenge
XX
XX Sequence 66 AA;
XX
XX Query Match 100.0%; Score 87; DB 5; Length 66;
XX Best Local Similarity 100.0%; Pred. No. 7.7e-07;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX OY 1 RSYQHDLRAYGFWR 15
XX Db 16 RSYQHDLRAYGFWR 30
XX
XX RESULT 4
XX AAE19244
XX ID AAE19244 standard; peptide; 15 AA.
XX
XX AC AAE19244;
XX
XX 21-MAY-2002 (first entry)
XX
XX Streptococcus pneumoniae PsaA immunogenic peptide #6.
XX Multiple antigenic peptide; MAP; immunogenic; immunity; infection;
XX pneumococcal surface adhesin protein A; PsaA; antibacterial.
XX
XX Streptococcus pneumoniae.
XX
XX WO200204497-A2.
```


CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ

SQ Sequence 554 AA;

Query Match 54.0%; Score 47; DB 4; Length 554;
 Best Local Similarity 53.8%; Pred. No. 31;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YQHDLRAYGFWRLL 15
 |:|:|:|:|:|:|:
 Db 7 YKHSUAYGAWAI 19

RESULT 7

ABG19731
 ID ABG19731 standard; protein; 554 AA.

XX AC ABG19731;
 XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #19722.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.

OS WO200175067-A2.
 FN 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US008631.
 XX 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 DR N-PSDB; AAS83918.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX Claim 20; SEQ ID NO 50090; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping.
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ

SQ Sequence 554 AA;

Query Match 54.0%; Score 47; DB 4; Length 554;
 Best Local Similarity 53.8%; Pred. No. 31;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YQHDLRAYGFWRLL 15
 |:|:|:|:|:|:|:
 Db 7 YKHSUAYGAWAI 19

RESULT 8

ABG03083
 ID ABG03083 standard; protein; 1007 AA.

XX AC ABG03083;
 XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #3074.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.

OS WO200175067-A2.
 FN 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US008631.
 XX 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 DR N-PSDB; AAS67270.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX Claim 20; SEQ ID NO 33442; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping.
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is

CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (ii) and its binding partners are useful in medical imaging
 CC of sites expressing (ii). (i) and (ii) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX
 SQ Sequence 1007 AA;

Query Match 54.0%; Score 47; DB 4; Length 1007;
 Best Local Similarity 53.8%; Pred. No. 58;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YQHDLRAYGFWR 15
 Db 7 YKSLHAYGAW 19
 ||| ||| |

RESULT 9
 ABP79868
 ID ABP79868 standard; protein; 741 AA.
 XX
 AC ABP79868;
 XX
 DT 07-MAR-2003 (first entry)
 XX
 DE N. gonorrhoeae amino acid sequence SEQ ID 6266.
 XX
 KW Antibacterial; infection; vaccine; gene therapy.
 XX
 OS Neisseria gonorrhoeae.
 XX
 PN WO200279243-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 12-FEB-2002; 2002WO-IB002069.
 XX
 PR 12-FEB-2001; 2001GB-00003424.
 XX
 PA (CHIR-) CHIRON SPA.
 XX
 PI Fontana MR, Pizza M, Massignani V, Monaci E;
 XX
 DR WPI; 2003-058415/05.
 DR N-PSDB; AB240838.
 XX
 PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
 PT medicament for treating or preventing N. gonorrhoeae infection.
 XX
 PS Disclosure; Page 645; 815pp; English.
 XX
 CC The present invention relates to proteins from Neisseria gonorrhoeae.
 CC Also disclosed are the nucleic acid molecules encoding the proteins and
 CC antibodies that specifically bind to the proteins. The composition
 CC comprising the protein, nucleic acid or antibody is useful for the
 CC manufacture of a medicament for treating or preventing N. gonorrhoeae
 CC infection, this may be in the form of a vaccine or gene therapy.
 CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
 CC molecules of the invention
 XX
 SQ Sequence 741 AA;

Query Match 51.7%; Score 45; DB 6; Length 741;
 Best Local Similarity 58.3%; Pred. No. 91;

CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (ii) and its binding partners are useful in medical imaging
 CC of sites expressing (ii). (i) and (ii) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX
 SQ Sequence 1007 AA;

Query Match 51.7%; Score 45; DB 7; Length 903;
 Best Local Similarity 53.8%; Pred. No. 11e+02;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 QHDLRAYGFWR 15
 Db 448 QHDVEAGGIWRM 459
 ||| ||| |

RESULT 10
 ADF03901
 ID ADF03901 standard; protein; 903 AA.
 XX
 AC ADF03901;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DE Bacterial polypeptide #14.
 XX
 KW Proteus mirabilis infection; bacterial infection; antibacterial;
 KW immunostimulant.
 XX
 OS Proteus mirabilis.
 XX
 PN US6605709-B1.
 XX
 PD 12-AUG-2003.
 XX
 PF 05-APR-2000; 2000US-00543681.
 XX
 PR 09-APR-1999; 99US-0128706P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton GL;
 XX
 DR WPI; 2003-895291/82.
 DR N-PSDB; ADE99728.
 XX
 PT New Proteus mirabilis polypeptides and polynucleotides, useful as
 PT reagents for diagnosis of bacterial disease, as components of
 PT antibacterial vaccines, as targets for antibacterial drugs, or as
 PT biocontrol agents for plants.
 XX
 PS Disclosure; SEQ ID NO 4186; 870pp; English.
 XX
 CC The invention relates to new Proteus mirabilis polypeptides and
 CC polynucleotides. The invention also relates to antibodies against the
 CC polypeptides, methods for producing the polypeptides, a method of
 CC generating vaccines for immunising an individual against P. mirabilis, a
 CC method for evaluating a compound for the ability to bind a P. mirabilis
 CC polypeptide and a method for screening test compounds for anti-bacterial
 CC activity. The polypeptides and polynucleotides are useful as molecular
 CC targets for diagnosing, preventing and treating pathological conditions
 CC resulting from bacterial infection, as reagents for diagnosis of
 CC bacterial diseases, as components of antibacterial vaccines, as targets
 CC for antibacterial drugs or as bio-control agents for plants. This
 CC sequence represents a Proteus mirabilis polypeptide of the invention.
 XX
 SQ Sequence 903 AA;

Query Match 51.7%; Score 45; DB 7; Length 903;
 Best Local Similarity 53.8%; Pred. No. 1.1e+02;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 SYQHDLRAYGFWR 14
 Db 110 NYRDLRLCYFWR 122
 ||| ||| |

RESULT 11
 ABP02252
 ID ABP02252 standard; protein; 118 AA.
 XX
 AC ABP02252;

Qy

KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.

XX Homo sapiens.

XX WO200157190-A2.

XX PD 09-AUG-2001.

XX PF 05-FEB-2001; 2001WO-US004098.

XX PR 03-FEB-2000; 2000US-00496914.

XX PR 27-APR-2000; 2000US-00560875.

XX PR 20-JUN-2000; 2000US-00598075.

XX PR 19-JUL-2000; 2000US-00620325.

XX PR 01-SEP-2000; 2000US-00654936.

XX PR 15-SEP-2000; 2000US-00663561.

XX PR 20-OCT-2000; 2000US-00693325.

XX PR 30-NOV-2000; 2000US-00728422.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

XX PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

XX PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX DR WPI; 2001-476283/51.

XX DR N-PSDB; AAK51596.

XX PT Nucleic acids encoding polypeptides with cytokine-like activities, useful

XX PT in diagnosis and gene therapy.

XX PS Claim 20; Page 3356-3359; 6221pp; English.

XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the

XX CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to

XX CC cytokine, cell proliferation or cell differentiation or which may induce

XX CC production of other cytokines in other cell populations. The

XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

XX CC peptide therapy. The polypeptides have various cytokine-like activities,

XX CC e.g. stem cell growth factor activity, haematopoiesis regulating

XX CC activity, tissue growth factor activity, immunomodulatory activity and

XX CC activin/inhibin activity and may be useful in the diagnosis and/or

XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

XX CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111

XX CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the

XX CC sequence listing were missing at the time of publication

XX SQ Sequence 1262 AA;

Query Match 50.6%; Score 44; DB 4; Length 1262;

Best Local Similarity 58.3%; Pred. No. 2.3e+02;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

XX OS Homo sapiens.

XX PN WO200157190-A2.

XX XX PD 09-AUG-2001.

XX XX PF 05-FEB-2001; 2001WO-US004098.

XX XX PR 03-FEB-2000; 2000US-00496914.

XX XX PR 27-APR-2000; 2000US-00560875.

XX XX PR 20-JUN-2000; 2000US-00598075.

XX XX PR 19-JUL-2000; 2000US-00620325.

XX XX PR 01-SEP-2000; 2000US-00654936.

XX XX PR 15-SEP-2000; 2000US-00663561.

XX XX PR 20-OCT-2000; 2000US-00693325.

XX XX PR 30-NOV-2000; 2000US-00728422.

XX XX PA (HYSE-) HYSEQ INC.

XX XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

XX XX PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

XX XX PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX XX DR WPI; 2001-476283/51.

XX XX DR N-PSDB; AAK52580.

XX PT Nucleic acids encoding polypeptides with cytokine-like activities, useful

XX PT in diagnosis and gene therapy.

XX PS Claim 20; Page 250-251; 6221pp; English.

XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the

XX CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to

XX CC cytokine, cell proliferation or cell differentiation or which may induce

XX CC production of other cytokines in other cell populations. The

XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

XX CC peptide therapy. The polypeptides have various cytokine-like activities,

XX CC e.g. stem cell growth factor activity, haematopoiesis regulating

XX CC activity, tissue growth factor activity, immunomodulatory activity and

XX CC activin/inhibin activity and may be useful in the diagnosis and/or

XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

XX CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111

XX CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the

XX CC sequence listing were missing at the time of publication

XX SQ Sequence 1289 AA;

Query Match 50.6%; Score 44; DB 4; Length 1289;

Best Local Similarity 58.3%; Pred. No. 2.4e+02;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

XX OS Homo sapiens.

XX PN WO200157190-A2.

XX XX PD 09-AUG-2001.

XX XX PF 05-FEB-2001; 2001WO-US004098.

XX XX PR 03-FEB-2000; 2000US-00496914.

XX XX PR 27-APR-2000; 2000US-00560875.

XX XX PR 20-JUN-2000; 2000US-00598075.

XX XX PR 19-JUL-2000; 2000US-00620325.

XX XX PR 01-SEP-2000; 2000US-00654936.

XX XX PR 15-SEP-2000; 2000US-00663561.

XX XX PR 20-OCT-2000; 2000US-00693325.

XX XX PR 30-NOV-2000; 2000US-00728422.

XX XX PA (HYSE-) HYSEQ INC.

XX XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

XX XX PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

XX XX PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX XX DR WPI; 2001-476283/51.

XX XX DR N-PSDB; AAK52580.

XX PT Nucleic acids encoding polypeptides with cytokine-like activities, useful

XX PT in diagnosis and gene therapy.

XX PS Claim 20; Page 250-251; 6221pp; English.

XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the

XX CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to

XX CC cytokine, cell proliferation or cell differentiation or which may induce

XX CC production of other cytokines in other cell populations. The

XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

XX CC peptide therapy. The polypeptides have various cytokine-like activities,

XX CC e.g. stem cell growth factor activity, haematopoiesis regulating

XX CC activity, tissue growth factor activity, immunomodulatory activity and

XX CC activin/inhibin activity and may be useful in the diagnosis and/or

XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

XX CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111

XX CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the

XX CC sequence listing were missing at the time of publication

XX SQ Sequence 1289 AA;

Query Match 50.6%; Score 44; DB 4; Length 1289;

Best Local Similarity 58.3%; Pred. No. 2.4e+02;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

XX OS Homo sapiens.

XX PN WO200157190-A2.

XX XX PD 09-AUG-2001.

XX XX PF 05-FEB-2001; 2001WO-US004098.

XX XX PR 03-FEB-2000; 2000US-00496914.

XX XX PR 27-APR-2000; 2000US-00560875.

XX XX PR 20-JUN-2000; 2000US-00598075.

XX XX PR 19-JUL-2000; 2000US-00620325.

XX XX PR 01-SEP-2000; 2000US-00654936.

XX XX PR 15-SEP-2000; 2000US-00663561.

XX XX PR 20-OCT-2000; 2000US-00693325.

XX XX PR 30-NOV-2000; 2000US-00728422.

XX XX PA (HYSE-) HYSEQ INC.

XX XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

XX XX PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

XX XX PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX XX DR WPI; 2001-476283/51.

XX XX DR N-PSDB; AAK52580.

XX PT Nucleic acids encoding polypeptides with cytokine-like activities, useful

XX PT in diagnosis and gene therapy.

XX PS Claim 20; Page 250-251; 6221pp; English.

XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the

XX CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to

XX CC cytokine, cell proliferation or cell differentiation or which may induce

XX CC production of other cytokines in other cell populations. The

XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

XX CC peptide therapy. The polypeptides have various cytokine-like activities,

XX CC e.g. stem cell growth factor activity, haematopoiesis regulating

XX CC activity, tissue growth factor activity, immunomodulatory activity and

XX CC activin/inhibin activity and may be useful in the diagnosis and/or

XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

XX CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111

XX CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the

XX CC sequence listing were missing at the time of publication

XX SQ Sequence 1289 AA;

Query Match 50.6%; Score 44; DB 4; Length 1289;

Best Local Similarity 58.3%; Pred. No. 2.4e+02;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

XX OS Homo sapiens.

XX PN WO200157190-A2.

XX XX PD 09-AUG-2001.

XX XX PF 05-FEB-2001; 2001WO-US004098.

XX XX PR 03-FEB-2000; 2000US-00496914.

XX XX PR 27-APR-2000; 2000US-00560875.

XX XX PR 20-JUN-2000; 2000US-00598075.

XX XX PR 19-JUL-2000; 2000US-00620325.

XX XX PR 01-SEP-2000; 2000US-00654936.

XX XX PR 15-SEP-2000; 2000US-00663561.

XX XX PR 20-OCT-2000; 2000US-00693325.

XX XX PR 30-NOV-2000; 2000US-00728422.

XX XX PA (HYSE-) HYSEQ INC.

XX XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

XX XX PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

XX XX PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX XX DR WPI; 2001-476283/51.

XX XX DR N-PSDB; AAK52580.

XX PT Nucleic acids encoding polypeptides with cytokine-like activities, useful

XX PT in diagnosis and gene therapy.

XX PS Claim 20; Page 250-251; 6221pp; English.

XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the

XX CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to

XX CC cytokine, cell proliferation or cell differentiation or which may induce

XX CC production of other cytokines in other cell populations. The

XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

XX CC peptide therapy. The polypeptides have various cytokine-like activities,

XX CC e.g. stem cell growth factor activity, haematopoiesis regulating

XX CC activity, tissue growth factor activity, immunomodulatory activity and

XX CC activin/inhibin activity and may be useful in the diagnosis and/or

XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

XX CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111

XX CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the

XX CC sequence listing were missing at the time of publication

XX SQ Sequence 1289 AA;

Query Match 50.6%; Score 44; DB 4; Length 1289;

Best Local Similarity 58.3%; Pred. No. 2.4e+02;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

XX OS Homo sapiens.

XX PN WO200157190-A2.

XX XX PD 09-AUG-2001.

XX XX PF 05-FEB-2001; 2001WO-US004098.

XX XX PR 03-FEB-2000; 2000US-00496914.

XX XX PR 27-APR-2000; 2000US-00560875.

XX XX PR 20-JUN-2000; 2000US-00598075.

XX XX PR 19-JUL-2000; 2000US-00620325.

XX XX PR 01-SEP-2000; 2000US-00654936.

XX XX PR 15-SEP-2000; 2000US-00663561.

XX XX PR 20-OCT-2000; 2000US-00693325.

XX XX PR 30-NOV-2000; 2000US-00728422.

XX XX PA (HYSE-) HYSEQ INC.

XX XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

XX XX PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

XX XX PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX XX DR WPI; 2001-476283/51.

XX XX DR N-PSDB; AAK52580.

XX PT Nucleic acids encoding polypeptides with cytokine-like activities, useful

XX PT in diagnosis and gene therapy.

XX PS Claim 20; Page 250-251; 6

CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1891 AA;

Query Match 50.6%; Score 44; DB 4; Length 1891;
Best Local Similarity 53.8%; Pred. No. 3.5e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RSYQHDLRAYGFW 13
|:|:|:|:|:
Db 120 RAYRADLRLGFW 132

RESULT 20
ABP40987
ID ABP40987 standard; protein; 29 AA.

XX AC ABP40987;
XX 01-AUG-2002 (first entry)
XX Human retroviral HERV-7q gag peptide #5.
XX Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
XX multiple sclerosis.

XX OS Human endogenous retrovirus.

XX PN WO9967395-A1.

XX PD 29-DEC-1999.

XX PF 23-JUN-1999; 99WO-FR001513.

XX PR 23-JUN-1998; 98FR-00007920.

XX PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX PI Alliel PM, Perin J, Rieger F;

XX WPI; 2000-160587/14.

XX New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
XX for diagnosis, treatment and prevention of autoimmune and neurological
XX diseases.

XX Claim 22; Page 151-152; 225pp; French.

XX The present invention relates to new nucleic acid sequences of human
XX endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
XX Regulatory elements associated with HERV-7q may alter expression of other
XX genes (even remote genes) on the same chromosome, inducing immunological
XX and/or neurological changes (which may be pathological or protective/
XX curative). HERV-7q peptides can be used to improve efficiency of the

CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
CC sequences can be used in immunogenic or vaccinating compositions, for
CC protection against autoimmune diseases, particularly multiple sclerosis.
CC The peptides may also be used (by sequence comparison) to detect/identify
CC endogenous retroviruses that are abnormally expressed in cancer,
CC neuropathologies or other autoimmune diseases. The present sequence was
CC used to illustrate the invention
XX
XX Sequence 29 AA;

Query Match 49.4%; Score 43; DB 3; Length 29;
Best Local Similarity 53.8%; Pred. No. 7;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YQHDRLRAYGFWRL 15
|:|:|:|:|:
Db 6 FSHDPQASGLWRL 18

RESULT 21
ABP80192
ID ABP80192 standard; protein; 126 AA.

XX AC ABP80192;

XX 07-MAR-2003 (first entry)

XX N. gonorrhoeae amino acid sequence SEQ ID 6914.

XX Antibacterial; infection; vaccine; gene therapy.

XX OS Neisseria gonorrhoeae.

XX PN WO200279243-A2.

XX PD 10-OCT-2002.

XX 12-FEB-2002; 2002WO-IB002069.

XX 12-FEB-2001; 2001GB-00003424.

XX (CHIR-) CHIRON SPA.

XX Fontana MR, Pizsa M, Massignani V, Monaci E;

XX WPI; 2003-058415/05.

XX N-PSDB; ABZ41162.

XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX medicament for treating or preventing N. gonorrhoeae infection.

XX Disclosure; Page 689; 815pp; English.

XX The present invention relates to proteins from Neisseria gonorrhoeae.
XX Also disclosed are the nucleic acid molecules encoding the proteins and
XX antibodies that specifically bind to the proteins. The composition
XX comprising the protein, nucleic acid or antibody is useful for the
XX manufacture of a medicament for treating or preventing N. gonorrhoeae
XX infection, this may be in the form of a vaccine or gene therapy.

XX Sequences given in records ABP76736-ABP81046 represent nucleic acid
XX molecules of the invention
XX
XX Sequence 126 AA;

Query Match 49.4%; Score 43; DB 6; Length 126;

Best Local Similarity 70.0%; Pred. No. 32;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 DLRAYGFWRL 15
|:|:|:|:|:
Db 94 DLGAYGFWSM 103

RESULT 22

ABP40989
ID ABP40989 standard; protein; 360 AA.

AC ABP40989;

DT 01-AUG-2002 (first entry)

DE Human retroviral gag protein.

KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
multiple sclerosis.

OS Human endogenous retrovirus.

XX WO9967395-A1.

XX 29-DEC-1999.

XX 23-JUN-1999; 99WO-FR001513.

XX 23-JUN-1998; 98FR-00007920.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX Alliel PM, Perin J, Rieger F;

XX WPI; 2000-160587/14.

PT New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
for diagnosis, treatment and prevention of autoimmune and neurological
diseases.

PS Claim 22; Page 154-155; 225pp; French.

XX The present invention relates to new nucleic acid sequences of human
endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
Regulatory elements associated with HERV-7q may alter expression of other
genes (even remote genes) on the same chromosome, inducing immunological
and/or neurological changes (which may be pathological or protective/
curative). HERV-7q peptides can be used to improve efficiency of the
immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
sequences can be used in immunogenic or vaccinating compositions, for
protection against autoimmune diseases, particularly multiple sclerosis.
The peptides may also be used (by sequence comparison) to detect/identify
endogenous retroviruses that are abnormally expressed in cancer,
neuropathologies or other autoimmune diseases. The present sequence was
used to illustrate the invention

XX Sequence 360 AA;

Query Match 49.4%; Score 43; DB 3; Length 360;
Best Local Similarity 53.8%; Pred. No. 93;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YQHLRAYGVFWRL 15

DB 336 FSHDPQASGLWRL 348

RESULT 23

ABP40990
ID ABP40990 standard; protein; 420 AA.

AC ABP40990;

DT 01-AUG-2002 (first entry)

DE Human retroviral gag protein.

KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
multiple sclerosis.

XX

OS Human endogenous retrovirus.

XX WO9967395-A1.

XX 29-DEC-1999.

XX 23-JUN-1999; 99WO-FR001513.

XX 23-JUN-1998; 98FR-00007920.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX Alliel PM, Perin J, Rieger F;

XX WPI; 2000-160587/14.

XX New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
for diagnosis, treatment and prevention of autoimmune and neurological
diseases.

XX Claim 22; Fig 6; 225pp; French.

XX The present invention relates to new nucleic acid sequences of human
endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
Regulatory elements associated with HERV-7q may alter expression of other
genes (even remote genes) on the same chromosome, inducing immunological
and/or neurological changes (which may be pathological or protective/
curative). HERV-7q peptides can be used to improve efficiency of the
immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
sequences can be used in immunogenic or vaccinating compositions, for
protection against autoimmune diseases, particularly multiple sclerosis.
The peptides may also be used (by sequence comparison) to detect/identify
endogenous retroviruses that are abnormally expressed in cancer,
neuropathologies or other autoimmune diseases. The present sequence was
used to illustrate the invention

XX Sequence 420 AA;

Query Match 49.4%; Score 43; DB 3; Length 420;
Best Local Similarity 53.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YQHLRAYGVFWRL 15

DB 337 FSHDPQASGLWRL 349

RESULT 24

ABP78342
ID ABP78342 standard; protein; 433 AA.

XX AC ABP78342;

XX DT 07-MAR-2003 (first entry)

XX DE N. gonorrhoeae amino acid sequence SEQ ID 3214.

XX KW Antibacterial; infection; vaccine; gene therapy.

XX OS Neisseria gonorrhoeae.

XX XX WO200279243-A2.

XX PD 10-OCT-2002.

XX PF 12-FEB-2002; 2002WO-IB002069.

XX PR 12-FEB-2001; 2001GB-00003424.

XX PA (CHIR-) CHIRON SPA.

XX PI Fontana MR, Pizza M, Massignani V, Monaci E;

XX

CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
 CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of the
 CC invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the manufacture
 CC of medicaments for treating or preventing infection due to *Neisseria*
 CC bacteria (e.g. meningitis and septicemia), to detect the presence of
 CC *Neisseria* bacteria, or to raise antibodies. They may also be used to
 CC screen for agonists or antagonists, which may themselves have use as
 CC antibacterial agents. The polynucleotides of the invention may also be
 CC used in gene therapy protocols. (Updated on 12-SEP-2003 to standardise OS
 CC field)
 XX Sequence 459 AA;
 SQ

Query Match 49.4%; Score 43; DB 3; Length 459;
 Best Local Similarity 50.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 SYQHDRLRAYGF 13
 :||:|||||
 Db 416 AYRFDMGIYGF 427

RESULT 27
 ABU63641
 ID ABU63641 standard; protein; 465 AA.
 XX AC ABU63641;
 XX 23-OCT-2003 (revised)
 DT 30-SEP-2003 (first entry)
 XX BsmAI restriction endonuclease, BsmAIR.
 XX Restriction endonuclease; enzyme; AsiSI; DNA cleaving; methylation;
 KW molecular cloning; BsmAIR.
 XX *Geobacillus stearothermophilus*.
 OS
 XX US2003104388-A1.
 XX 05-JUN-2003.
 XX 20-SEP-2001; 2001US-00957005.
 XX 20-SEP-2001; 2001US-00957005.
 XX (NEWE) NEW ENGLAND BIOLABS INC.
 PA Zhu Z, Zhou J, Xu S;
 PI WPI: 2003-567952/53.
 DR N-PSDB; ACD28244.
 XX New DNA coding for the AsiSI restriction endonuclease or AsiSI methylase,
 PT useful for cleaving DNA molecules into small fragments for molecular
 PT cloning and gene characterization.
 XX Disclosure; Fig 3; 27pp; English.
 XX The invention relates to a new isolated DNA coding for the AsiSI
 CC restriction endonuclease or AsiSI methylase. The DNA is useful for
 CC cleaving DNA molecules into small fragments for molecular cloning and
 CC gene characterisation. The DNA is also useful for modifying DNA molecules
 CC via methylation. The present sequence represents the amino acid sequence
 CC of the BsmAI restriction endonuclease, BsmAIR. (Updated on 23-OCT-2003 to
 CC standardise OS field)
 XX Sequence 465 AA;
 SQ

Query Match 49.4%; Score 43; DB 6; Length 465;
 Best Local Similarity 53.8%; Pred. No. 1.2e+02;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 RSYQHDRLRAYGF 13
 :||:|||||
 Db 214 KSYTKDRAYEYW 226

RESULT 28
 ABU32026
 ID ABU32026 standard; protein; 466 AA.
 XX AC ABU32026;
 XX 19-JUN-2003 (first entry)
 XX Protein encoded by Prokaryotic essential gene #17553.
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 KW Klebsiella pneumoniae.
 OS
 XX WO200277183-A2.
 XX 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI: 2003-029926/02.
 DR N-PSDB; ACA35896.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 25; SEQ ID NO 59950; 1766pp; English.
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational

```

Query Match      49.4%; Score 43; DB 7; Length 484;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0
Qy      8 RAYGEFWRLL 15
Db      448 RRYGEFWRLL 455

```

| | | | |
|----|----|------|------|
| DT | XX | Phob | 20-N |
| DE | XX | Phob | |
| KW | KW | Anti | |
| KW | KW | dete | |
| KW | KW | anti | |
| KW | KW | whoo | |
| XX | XX | Phob | |
| OS | OS | Phob | |

XX WO200294867-A2.
 XX
 PD 28-NOV-2002.
 XX
 XX 07-FEB-2002; 2002WO-IB003040.
 XX
 XX 07-FEB-2001; 2001FR-00001659.
 XX
 XX (INSP) INST PASTEUR.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 XX Duchaud E, Taourit S, Glaeer P, Frangeul L, Kunst F, Danchin A,
 PI Buchrieser C;
 XX
 XX WPI; 2003-148459/14.
 XX
 XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 XX
 XX Claim 2; SEQ ID NO 1942; 1205pp; French.
 XX
 CC The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterials useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plaque and whooping cough). This
 CC sequence represents one of the isolated P. luminescens proteins
 XX
 XX Sequence 280 AA;
 SQ
 Query Match 48.3%; Score 42; DB 6; Length 280;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 5 HDLRAYGFW 13
 | | | | |
 Db 171 HSRRSYGFW 179
 RESULT 32
 ABB67974
 ID ABB67974 standard; protein; 652 AA.
 XX
 AC ABB67974;
 XX
 XX 26-MAR-2002 (first entry)
 DT
 XX Drosophila melanogaster polypeptide SEQ ID NO 30714.
 DE
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 XX Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD

XX 23-MAR-2001; 2001WO-US009231.
 PF
 XX 23-MAR-2000; 2000US-0191637P.
 PR
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 XX
 DR N-PSDB; ABL12077.
 DR
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 XX Disclosure; SEQ ID NO 30714; 21pp + Sequence Listing; English.
 PS
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABU01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 652 AA;
 SQ
 Query Match 48.3%; Score 42; DB 4; Length 652;
 Best Local Similarity 53.8%; Pred. No. 2.5e+02;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 RSYQHDRLRAYGFW 13
 | | | | |
 Db 557 RSFQYSLTDYHFW 569
 RESULT 33
 AAY34536
 ID AAY34536 standard; protein; 951 AA.
 XX
 AC AAY34536;
 XX
 XX 27-AUG-2003 (revised)
 DT
 DT 20-MAR-2003 (revised)
 DT 25-AUG-1999 (first entry)
 XX
 XX Porphyromonas gingivalis protein PG67.
 DE
 XX Porphyromonas gingivalis; PG; periodontal disease; gingivitis; vaccine;
 KW antigenic.
 KW
 XX Porphyromonas gingivalis.
 OS
 XX WO9929870-A1.
 PN
 XX 17-JUN-1999.
 PD
 XX 10-DEC-1998; 98WO-AU001023.
 PF
 XX 10-DEC-1997; 97AU-00000839.
 PR
 PR 31-DEC-1997; 97AU-00001182.
 PR 30-JAN-1998; 98AU-00001546.
 PR 10-MAR-1998; 98AU-00002264.
 PR 09-APR-1998; 98AU-00002911.
 PR 23-APR-1998; 98AU-00003128.
 PR 05-MAY-1998; 98AU-00003338.
 PR 22-MAY-1998; 98AU-00003654.
 PR 29-JUL-1998; 98AU-00004917.
 PR

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 111 AA;

Query Match 47.1%; Score 41; DB 4; Length 111;
 Best Local Similarity 38.5%; Pred. No. 60;
 Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 RSYQHDRLRAYGFW 13
 | : | : | : |
 Db 51 RGFRRHDIRGFAQW 63

RESULT 36
 ABG25967
 ID ABG25967 standard; protein; 111 AA.
 AC ABG25967;
 XX
 XX
 DT 18-FEB-2002 (first entry)
 DE Novel human diagnostic protein #25958.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 XX
 PN WO200175067-A2.
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Drmanac RT, Liu C, Tang YT;
 PI
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS90154.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 56326; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 111 AA;

Query Match 47.1%; Score 41; DB 4; Length 111;
 Best Local Similarity 38.5%; Pred. No. 60;
 Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 RSYQHDRLRAYGFW 13
 | : | : | : |
 Db 51 RGFRRHDIRGFAQW 63

RESULT 37
 ADF05142
 ID ADF05142 standard; protein; 144 AA.
 AC ADF05142;
 XX
 XX
 DT 12-FEB-2004 (first entry)
 DE Bacterial polypeptide #1255.
 XX
 KW Proteus mirabilis infection; bacterial infection; antibacterial;
 KW immunostimulant.
 XX
 OS Proteus mirabilis.
 XX
 PN US6605709-B1.
 PD 12-AUG-2003.
 XX
 PF 05-APR-2000; 2000US-00543681.
 XX
 PR 09-APR-1999; 99US-0128706P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton GL;
 XX
 DR WPI; 2003-895291/82.
 DR N-PSDB; ADF00970.
 XX
 XX New Proteus mirabilis polypeptides and polynucleotides, useful as
 PT reagents for diagnosis of bacterial disease, as components of
 PT antibacterial vaccines, as targets for antibacterial drugs, or as
 PT biocontrol agents for plants.
 XX
 PS Disclosure; SEQ ID NO 5427; 870pp; English.
 XX
 CC The invention relates to new Proteus mirabilis polypeptides and
 CC polynucleotides. The invention also relates to antibodies against the
 CC polypeptides, methods for producing the polypeptides, a method of
 CC generating vaccines for immunising an individual against P. mirabilis, a
 CC method for evaluating a compound for the ability to bind a P. mirabilis
 CC polypeptide and a method for screening test compounds for anti-bacterial
 CC activity. The polypeptides and polynucleotides are useful as molecular

CC targets for diagnosing, preventing and treating pathological conditions
CC resulting from bacterial infection, as reagents for diagnosis of
CC bacterial diseases, as components of antibacterial vaccines, as targets
CC for antibacterial drugs or as bio-control agents for plants. This
CC sequence represents a *Proteus mirabilis* polypeptide of the invention.
XX
SQ Sequence 144 AA;

Query Match 47.1%; Score 41; DB 7; Length 144;
Best Local Similarity 87.5%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 YQHDLRAY 10
DB 56 YQHDLRAY 63
|||||

RESULT 38
ABO76461
ID ABO76461 standard; protein; 227 AA.
XX ABO76461;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #8636.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
XX WPI; 2003-615309/58.
DR N-PSDB; ABD10032.
XX

PT Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide, of
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 25207; 455pp; English.
XX
CC The invention relates to *Pseudomonas aeruginosa* polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-*P. aeruginosa* drugs, as templates for recombinant
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
CC infection, and in detection of *P. aeruginosa* sequences or other sequences
CC of *Pseudomonas* species using biochip technology. Sequences ABO7826-
CC ABO84396 represent *P. aeruginosa* polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX

SQ Sequence 227 AA;

Query Match 47.1%; Score 41; DB 7; Length 227;
Best Local Similarity 87.5%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Best Local Similarity 63.6%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 5 HDLRAYGFWRLL 15
DB 14 HDAPAYGLWLL 24
|||||

RESULT 39
ADI27244
ID ADI27244 standard; protein; 497 AA.
XX ADI27244;
XX
DT 22-APR-2004 (first entry)
XX
DE B. napus BnAOX-1 amino acid sequence, SEQ ID 5.
XX
KW Plant; protectant; amine oxidase; stress response;
KW stress-related polypeptide; AOXSRP; environmental stress; high salinity;
KW drought; low temperature; lodging.
XX
OS Brassica napus.
XX
PN WO2003039243-A2.
XX
PD 15-MAY-2003.
XX
PF 12-NOV-2002; 2002WO-US036507.
XX
PR 09-NOV-2001; 2001US-0345969P.
XX
PA (BADI) BASF PLANT SCI GMBH.
XX
PI Henkes S, Chen R, Da Costa E SilvaO, Van Thielen N;
XX
XX WPI; 2003-430616/40.
DR N-PSDB; ADI27243.
XX

PT New amine oxidase stress-related polypeptides and nucleic acids encoding
PT them, useful for increasing a plant's tolerance to environmental stress,
PT including high salinity, drought, low temperature and lodging.
XX
PS Claim 1; Page 71; 77pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising a
CC polynucleotide encoding a polypeptide that functions as a modulator of a
CC plant stress response. Also disclosed as new are amino oxidase stress-
CC related polypeptide (AOXSRP) chimeric or fusion polypeptides comprising
CC an AOXSRP fused to a non-AOXSRP polypeptide, and antibodies that
CC specifically bind to an AOXSRP. The amine oxidase stress-related
CC polypeptides are useful for increasing a plant's tolerance to
CC environmental stress, including high salinity, drought, low temperature
CC and lodging, or for transforming plants tolerant to these environmental
CC stresses. The nucleic acid molecules, polypeptides, peptide homologs,
CC fusion polypeptides, primers, vectors and host cells are useful in the
CC identification of Physcomitrella patens, Brassica napus and other related
CC organisms. They are useful for the mapping of genomes of organisms
CC related to *P. patens* or *B. napus*, the identification and localisation of
CC *P. patens* or *B. napus* sequences of interest, the determination of AOXSRP
CC regions required for function, the modulation of AOXSRP activity, the
CC modulation of expression of an AOXSRP coding nucleic acid, the modulation
CC of metabolism or one or more cell functions, the modulation of
CC transmembrane transport of one or more compounds, or the modulation of
CC stress resistance. The current sequence represents the *B. napus* BnAOX-1
CC amino acid sequence.
XX

SQ Sequence 497 AA;

Query Match 47.1%; Score 41; DB 7; Length 497;
Best Local Similarity 54.5%; Pred. No. 2.8e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 YQHDLRAYGFW 13
| | | | : | | :
Db 115 YQHDLESYGLY 125

RESULT 40

ADG98169
ID ADG98169 standard; protein; 544 AA.

XX AC ADG98169;

XX DT 11-MAR-2004 (first entry)

XX DE BsaI restriction endonuclease (bsaIR) polypeptide.

XX BsaI restriction endonuclease; endonuclease; methylase;
KW restriction enzyme; enzyme purification; bsaIR; enzyme.

XX OS Geobacillus stearothermophilus.

XX PN US2003186363-A1.

XX PD 02-OCT-2003.

XX PF 26-MAR-2002; 2002US-00106275.

XX PR 26-MAR-2002; 2002US-00106275.

XX PA (NEWE) NEW ENGLAND BIOLABS INC.

XX PI Zhu Z, Xu S;

XX DR WPI; 2004-032006/03.

DR N-PSDB; ADG98168.

XX PT New isolated DNA obtainable from Bacillus stearothermophilus used in
PT vectors for transforming a host cell and producing recombinant BsaI
PT restriction endonuclease.

XX PS Disclosure; SEQ ID NO 6; 29pp; English.

XX CC The invention relates to DNA encoding the BsaI restriction endonuclease,
CC where the DNA is obtainable from Bacillus stearothermophilus 6-55. The
CC invention also relates to a recombinant DNA vector into which a DNA
CC segment encoding the BsaI restriction endonuclease gene has been
CC inserted, and a method for producing recombinant BsaI restriction
CC endonuclease comprising culturing a host cell transformed with the vector
CC under conditions suitable for expression of endonuclease and methylase.
CC The invention is used to produce large quantities of restriction enzymes,
CC thus simplifying enzyme purification. This sequence represents the bsaIR
CC restriction endonuclease polypeptide of the invention.

XX SQ Sequence 544 AA;

Query Match 47.1%; Score 41; DB 8; Length 544;
Best Local Similarity 53.8%; Pred. No. 3.1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 RSYQHDLRAYGFW 13

Db 229 RSYTDRRAFEYW 241

Search completed: March 18, 2005, 18:27:44
Job time : 162 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 18:22:41 ; Search time 42 Seconds
(without alignments)
26.660 Million cell updates/sec

Title: US-09-623-038-6

Perfect score: 87

Sequence: 1 RSYQHDRLRAYGFWR 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgm2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgm2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgm2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgm2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgm2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgm2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 45 | 51.7 | 903 | 4 | US-09-543-681A-4186 |
| 2 | 43 | 49.4 | 465 | 4 | US-09-957-005-11 |
| 3 | 43 | 49.4 | 484 | 4 | US-09-489-039A-9895 |
| 4 | 43 | 49.4 | 530 | 4 | US-09-966-997-13 |
| 5 | 41 | 47.1 | 92 | 4 | US-09-270-767-38503 |
| 6 | 41 | 47.1 | 92 | 4 | US-09-270-767-53720 |
| 7 | 41 | 47.1 | 144 | 4 | US-09-543-681A-5427 |
| 8 | 41 | 47.1 | 227 | 4 | US-09-252-991A-25207 |
| 9 | 41 | 47.1 | 544 | 4 | US-10-106-275-6 |
| 10 | 40 | 46.0 | 151 | 4 | US-09-765-815-13 |
| 11 | 40 | 46.0 | 334 | 4 | US-09-287-599A-2 |
| 12 | 40 | 46.0 | 920 | 4 | US-09-934-868-52 |
| 13 | 39 | 44.8 | 239 | 4 | US-09-248-796A-22464 |
| 14 | 39 | 44.8 | 265 | 1 | US-08-413-803-27 |
| 15 | 39 | 44.8 | 265 | 5 | PCT-US95-03776-29 |
| 16 | 39 | 44.8 | 286 | 1 | US-08-321-488A-27 |
| 17 | 39 | 44.8 | 433 | 4 | US-09-711-164-412 |
| 18 | 39 | 44.8 | 504 | 4 | US-09-248-796A-17765 |
| 19 | 39 | 44.8 | 2104 | 2 | US-08-808-793-4 |
| 20 | 39 | 44.8 | 2104 | 3 | US-08-772-512A-4 |
| 21 | 39 | 44.8 | 2104 | 4 | US-09-428-371-4 |
| 22 | 39 | 44.8 | 2105 | 2 | US-08-808-793-3 |
| 23 | 39 | 44.8 | 2105 | 3 | US-08-772-512A-3 |
| 24 | 39 | 44.8 | 2105 | 4 | US-09-428-371-3 |
| 25 | 38 | 43.7 | 66 | 3 | US-08-936-165A-376 |
| 26 | 38 | 43.7 | 109 | 4 | US-09-328-352-7204 |
| 27 | 38 | 43.7 | 148 | 3 | US-08-946-329A-65 |

ALIGNMENTS

RESULT 1

US-09-543-681A-4186
; Sequence 4186, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4186
; LENGTH: 903
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4186

Query Match 51.7%; Score 45; DB 4; Length 903;
Best Local Similarity 53.8%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 SYQHDRLRAYGFWR 14
Db 110 NYRHDRLGYFWR 122

RESULT 2

US-09-957-005-11
; Sequence 11, Application US/09957005
; Patent No. 6596524
; GENERAL INFORMATION:
; APPLICANT: Zhu, Zhenyu
; APPLICANT: Zhou, Jing
; APPLICANT: Xu, Shuang-yong
; TITLE OF INVENTION: Method For Cloning And Expression of BsmAI Restriction
; FILE REFERENCE: NEB-192
; CURRENT APPLICATION NUMBER: US/09/957,005
; CURRENT FILING DATE: 2001-09-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus A664
US-09-957-005-11

28 38 43.7 181 4 US-09-107-532A-4183 Sequence 4183, Ap
29 38 43.7 320 1 US-08-362-739-2 Sequence 2, Appli
30 38 43.7 320 4 US-08-914-350A-2 Sequence 2, Appli
31 38 43.7 557 4 US-09-902-540-16548 Sequence 16548, A
32 38 43.7 681 4 US-09-252-991A-28831 Sequence 28831, A
33 38 43.7 827 4 US-09-902-540-13652 Sequence 13652, A
34 38 43.7 1111 1 US-08-317-450B-15 Sequence 15, Appl
35 38 43.7 1111 3 US-08-800-593-15 Sequence 15, Appl
36 38 43.7 1171 1 US-08-445-135-1 Sequence 1, Appli
37 38 43.7 1171 4 US-09-560-385A-36 Sequence 36, Appl
38 38 43.7 1172 4 US-09-560-385A-32 Sequence 32, Appl
39 38 43.7 1172 4 US-09-560-385A-34 Sequence 34, Appl
40 38 43.7 1192 1 US-08-317-450B-13 Sequence 13, Appl
41 38 43.7 1193 3 US-08-800-593-13 Sequence 13, Appl
42 38 43.7 1193 4 US-09-560-385A-26 Sequence 26, Appl
43 38 43.7 1193 4 US-09-560-385A-30 Sequence 30, Appl
44 38 43.7 1193 4 US-09-560-385A-30 Sequence 2, Appli
45 37 42.5 16 4 US-09-344-783C-2

Query Match 49.4%; Score 43; DB 4; Length 465;
Best Local Similarity 53.8%; Pred. No. 21;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RSYQHDRLRAYGFW 13
Db 214 KSYTDRLRAYEYX 226

RESULT 3

US-09-489-039A-9895
; Sequence 9895, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9895
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9895

Query Match 49.4%; Score 43; DB 4; Length 484;
Best Local Similarity 87.5%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 8 RAYGFWRL 15
Db 448 RRYGFWRL 455

RESULT 4

US-09-966-997-13
; Sequence 13, Application US/09966997
; Patent No. 6764843
; GENERAL INFORMATION:
; APPLICANT: Xu, Shuang-yong
; APPLICANT: Dore, Andrew
; APPLICANT: Hume, Adam
; APPLICANT: Pelletier, John
; APPLICANT: Zhou, Jing
; TITLE OF INVENTION: Method For Cloning And Expression of BamBI Restriction
; TITLE OF INVENTION: Endonuclease And BamBI Methylase In E. coli And
; FILE REFERENCE: NEB-191
; CURRENT APPLICATION NUMBER: US/09/966,997
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus B61
US-09-966-997-13

Query Match 49.4%; Score 43; DB 4; Length 530;
Best Local Similarity 53.8%; Pred. No. 24;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RSYQHDRLRAYGFW 13
Db 226 RLYNHDRAPFWX 238

RESULT 5

US-09-270-767-38503

; Sequence 38503, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38503
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-38503

Query Match 47.1%; Score 41; DB 4; Length 92;
Best Local Similarity 87.5%; Pred. No. 8.1;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 DLRAYGFW 13
Db 45 DLRAYPFW 52

RESULT 6

US-09-270-767-53720
; Sequence 53720, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 53720
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-53720

Query Match 47.1%; Score 41; DB 4; Length 92;
Best Local Similarity 87.5%; Pred. No. 8.1;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 DLRAYGFW 13
Db 45 DLRAYPFW 52

RESULT 7

US-09-543-681A-5427
; Sequence 5427, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5427
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5427

Query Match 47.1%; Score 41; DB 4; Length 144;

```
Best Local Similarity 87.5%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 YQHLRAY 10
   |||||
Db 56 YQHLREY 63

RESULT 8
US-09-252-991A-25207
; Sequence 25207, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25207
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25207

Query Match 47.1%; Score 41; DB 4; Length 227;
Best Local Similarity 63.6%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 HDLRAYFWRL 15
   |||||
Db 14 HDPAYGLWLL 24

RESULT 9
US-10-106-275-6
; Sequence 6, Application US/10106275
; Patent No. 6723546
; GENERAL INFORMATION:
; APPLICANT: Xu, Shuang-yong
; TITLE OF INVENTION: Method For Cloning And Expression of Bsal Restriction Endonuclease
; FILE REFERENCE: NEB-202
; CURRENT APPLICATION NUMBER: US/10/106,275
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus 6-55
US-10-106-275-6

Query Match 47.1%; Score 41; DB 4; Length 544;
Best Local Similarity 53.8%; Pred. No. 55;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 RSYQHLRAYGF 13
   |||||
Db 229 RSYTDRAPFW 241

RESULT 10
US-09-765-815-13
; Sequence 13, Application US/09765815
; Patent No. 6673586
; GENERAL INFORMATION:
; APPLICANT: Balk, Steven
; TITLE OF INVENTION: Interacting Protein Kinase
; FILE REFERENCE: 01948/068002
; CURRENT APPLICATION NUMBER: US/09/765,815
; CURRENT FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/176,859
; PRIOR FILING DATE: 2000-01-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-765-815-13

Query Match 46.0%; Score 40; DB 4; Length 151;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 RSYQHLRAYGF 12
   |||||
Db 52 RDYQHEVEMYNF 63

RESULT 11
US-09-287-599A-2
; Sequence 2, Application US/09287599A
; Patent No. 6602712
; GENERAL INFORMATION:
; APPLICANT: Handelsman, Jo
; TITLE OF INVENTION: Enterotoxin-Deficient Bacillus
; FILE REFERENCE: 960296.95327
; CURRENT APPLICATION NUMBER: US/09/287,599A
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/080943
; PRIOR FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Bacillus
; OTHER INFORMATION: cereus hblA, having 123 bp in-frame deletion
; OTHER INFORMATION: spanning nucleotides 568-690
US-09-287-599A-2

Query Match 46.0%; Score 40; DB 4; Length 334;
Best Local Similarity 63.6%; Pred. No. 48;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SYQHLRAYGF 12
   |||||
Db 181 SIGHDVRAFGE 191

RESULT 12
US-09-934-868-52
; Sequence 52, Application US/09934868
; Patent No. 6689601
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odom, James M
; APPLICANT: Schenzle, Andreas J
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1596 US NA
; CURRENT APPLICATION NUMBER: US/09/934,868
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
```

NUMBER OF SEQ ID NOS: 81
SOFTWARE: Microsoft Office 97
SEQ ID NO 52
LENGTH: 920
TYPE: PRT
ORGANISM: Methylobionas 16a
FEATURE:
OTHER INFORMATION: Naga
US-09-934-868-52

Query Match 46.0%; Score 40; DB 4; Length 920;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YQHDRLRAYGFWR 14
| : : : : :
DB 727 YHVKGHGFWR 738

RESULT 13
US-09-248-796A-22464
Sequence 22464, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 22464
LENGTH: 239
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-22464

Query Match 44.8%; Score 39; DB 4; Length 239;
Best Local Similarity 58.3%; Pred. No. 50;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 YQHDRLRAYGFWR 14
| : : : : :
DB 2 YVHDLRKFFFR 13

RESULT 14
US-08-413-803-27
Sequence 27, Application US/08413803
Patent No. 5766581
GENERAL INFORMATION:
APPLICANT: Bartley, Timothy D.
APPLICANT: Bogenberger, Jakob M.
APPLICANT: Besselman, Robert A.
APPLICANT: Hunt, Pamela
APPLICANT: Kinstler, Olaf B.
APPLICANT: Samal, Babru B.
TITLE OF INVENTION: METHODS FOR TREATING MAMMALS WITH
TITLE OF INVENTION: MONO-PEGYLATED PROTEINS THAT STIMULATE MEGAKARYOCYTE
GROWTH AND DIFFERENTIATION
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: AMGEN INC.
STREET: 1840 DeHavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: US
ZIP: 91320-1789
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/413,803
FILING DATE: 30-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,768
FILING DATE: 31-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/252,628
FILING DATE: 31-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/321,488
FILING DATE: 12-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/347,780
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Cook Ph.D. Robert R.
REGISTRATION NUMBER: 31,602
REFERENCE/DOCKET NUMBER: A-290D
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 265 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-413-803-27

Query Match 44.8%; Score 39; DB 1; Length 265;
Best Local Similarity 58.3%; Pred. No. 56;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 SYQHDRLRAYGFW 13
| : : : : :
DB 130 SFQHLRGKDFW 141

RESULT 15
PCT-US95-03776-29
Sequence 29, Application PC/TUS9503776
GENERAL INFORMATION:
APPLICANT: AMGEN INC.
TITLE OF INVENTION: Compositions and Methods for Stimulating
TITLE OF INVENTION: Megakaryocyte Growth and Differentiation
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 DeHavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/03776
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Cook, Robert R.
REFERENCE/DOCKET NUMBER: A-290-C
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 265 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
PCT-US95-03776-29

Query Match 44.8%; Score 39; DB 5; Length 265;
Best Local Similarity 58.3%; Pred. No. 56;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 SYQHDLRAYGF 13
|:|:|:|:|:|:|
Db 130 SFQHLRGKDFW 141

RESULT 16

US-08-321-488A-27
; Sequence 27, Application US/08321488A
; Patent No. 5795569

GENERAL INFORMATION:

; APPLICANT: Bartley, Timothy D.
; APPLICANT: Bogenberger, Jakob M.
; APPLICANT: Bosselman, Robert A.
; APPLICANT: Hunt, Pamela
; APPLICANT: Kinstler, Olaf B.
; APPLICANT: Samal, Babru B.
; TITLE OF INVENTION: MONO-PREGLYLATED PROTEINS THAT STIMULATE
; TITLE OF INVENTION: MEGAKARYOCYTE GROWTH AND DIFFERENTIATION
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESS: AMGEN INC.
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: US
; ZIP: 91320-1789

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/321,488A
; FILING DATE: 12-OCT-1994
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/252,628
; FILING DATE: 31-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,768
; FILING DATE: 31-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Cook, Robert R.

; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-290B
; INFORMATION FOR SEQ ID NO: 27:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 286 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-321-488A-27

Query Match 44.8%; Score 39; DB 1; Length 286;
Best Local Similarity 58.3%; Pred. No. 61;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 SYQHDLRAYGF 13
|:|:|:|:|:|:|
Db 151 SFQHLRGKDFW 162

RESULT 17

US-09-711-164-412
; Sequence 412, Application US/09711164
; Patent No. 6589738

GENERAL INFORMATION:

; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERE-
; FILE REFERENCE: ELITRA.008A
; CURRENT APPLICATION NUMBER: US/09/711,164
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-9
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 412
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-711-164-412

Query Match 44.8%; Score 39; DB 4; Length 433;
Best Local Similarity 46.2%; Pred. No. 95;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 YQHDLRAYGF 15
|:|:|:|:|:|:|
Db 410 FWHLRRLGLMHL 422

RESULT 18

US-09-248-796A-17765
; Sequence 17765, Application US/09248796A
; Patent No. 6747137

GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17765

; LENGTH: 504
; TYPE: PRT

; ORGANISM: Candida albicans
; FEATURE:

; NAME/KEY: UNSURE
; LOCATION: (481)

; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unk

US-09-248-796A-17765

Query Match 44.8%; Score 39; DB 4; Length 504;
Best Local Similarity 54.5%; Pred. No. 11e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 HDLRAYGF 15
|:|:|:|:|:|:|
Db 166 HDERDFWKL 176

RESULT 19

US-08-808-793-4
; Sequence 4, Application US/08808793
; Patent No. 5858713

GENERAL INFORMATION:

; APPLICANT: Soderlund, David M.
; APPLICANT: Ingles, Patricia J.
; TITLE OF INVENTION: CALCIUM PERMEABLE INSECT SODIUM CHANNELS
; TITLE OF INVENTION: AND USE THEREOF
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,793
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/034,361
FILING DATE: 24-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,649
FILING DATE: 01-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Bramean, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 19603/1062 (D-1906A)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1636
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2104 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-808-793-4

Query Match 44.8%; Score 39; DB 2; Length 2104;
Best Local Similarity 53.8%; Pred. No. 5.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YQHDRLRAYGFWRL 15
||| ||| : : :
Db 383 YQHVLAAGPWHM 395

RESULT 20
US-08-772-512A-4
; Sequence 4, Application US/08772512A
; Patent No. 6022705
; GENERAL INFORMATION:
; APPLICANT: Soderlund, David M.
; APPLICANT: Knipple, Douglas C.
; APPLICANT: Ingles, Patricia J.
; TITLE OF INVENTION: INSECT SODIUM CHANNELS FROM
; TITLE OF INVENTION: INSECTICIDE-SUSCEPTIBLE AND INSECTICIDE-RESISTANT HOUSE
; TITLE OF INVENTION: FLIES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/772,512A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/608,618
FILING DATE: 01-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bramean, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 19603/601 (CRFD-1657)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1636
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2104 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-772-512A-4

Query Match 44.8%; Score 39; DB 3; Length 2104;
Best Local Similarity 53.8%; Pred. No. 5.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YQHDRLRAYGFWRL 15
||| ||| : : :
Db 383 YQHVLAAGPWHM 395

RESULT 21
US-09-428-371-4
; Sequence 4, Application US/09428371
; Patent No. 6800435
; GENERAL INFORMATION:
; APPLICANT: Soderlund, David M.
; APPLICANT: Knipple, Douglas C.
; APPLICANT: Ingles, Patricia J.
; TITLE OF INVENTION: INSECT SODIUM CHANNELS FROM INSECTICIDE-SUSCEPTIBLE AND
; TITLE OF INVENTION: INSECTICIDE-RESISTANT HOUSE FLIES
; FILE REFERENCE: 19603/606
; CURRENT APPLICATION NUMBER: US/09/428,371
; CURRENT FILING DATE: 1999-10-28
; EARLIER APPLICATION NUMBER: 08/608,618
; EARLIER FILING DATE: 1996-03-01
; EARLIER APPLICATION NUMBER: 08/772,512
; EARLIER FILING DATE: 1996-12-24
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2104
; TYPE: PRT
; ORGANISM: Musca domestica
US-09-428-371-4

Query Match 44.8%; Score 39; DB 4; Length 2104;
Best Local Similarity 53.8%; Pred. No. 5.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YQHDRLRAYGFWRL 15
||| ||| : : :
Db 383 YQHVLAAGPWHM 395

RESULT 22
US-08-808-793-3
; Sequence 3, Application US/08808793
; Patent No. 5858713
; GENERAL INFORMATION:
; APPLICANT: Soderlund, David M.
; APPLICANT: Ingles, Patricia J.
; TITLE OF INVENTION: CALCIUM PERMEABLE INSECT SODIUM CHANNELS

;; TITLE OF INVENTION: AND USE THEROF
;; NUMBER OF SEQUENCES: 32
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
;; STREET: Clinton Square, P.O. Box 1051
;; CITY: Rochester
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 14603

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/808,793

;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 60/034,361
;; FILING DATE: 24-DEC-1996

;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Braman, Susan J.
;; REGISTRATION NUMBER: 34,103
;; REFERENCE/DOCKET NUMBER: 19603/1062 (D-1906A)
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 716-263-1636
;; TELEFAX: 716-263-1600

;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2105 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-808-793-3

Query Match 44.8%; Score 39; DB 2; Length 2105;
Best Local Similarity 53.8%; Pred. No. 5.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YOHDLRAYGFWRLL 15
|||:|:|:|:
Db 383 YOHVLOAAGPWHM 395

RESULT 23
US-08-772-512A-3
; Sequence 3, Application US/08772512A
; Patent No. 6022705

;; GENERAL INFORMATION:
;; APPLICANT: Soderlund, David M.
;; APPLICANT: Knipple, Douglas C.
;; APPLICANT: Ingles, Patricia J.

;; TITLE OF INVENTION: INSECT SODIUM CHANNELS FROM
;; TITLE OF INVENTION: INSECTICIDE-SUSCEPTIBLE AND INSECTICIDE-RESISTANT HOUSE
;; TITLE OF INVENTION: FLIES
;; NUMBER OF SEQUENCES: 19

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
;; STREET: P.O. Box 1051, Clinton Square
;; CITY: Rochester
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 14603

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/772,512A

;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/608,618
;; FILING DATE: 01-MAR-1996

;; ATTORNEY/AGENT INFORMATION:
;; NAME: Braman, Susan J.
;; REGISTRATION NUMBER: 34,103
;; REFERENCE/DOCKET NUMBER: 19603/601 (CRFD-1657)
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 716-263-1636
;; TELEFAX: 716-263-1600

;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2105 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-772-512A-3

Query Match 44.8%; Score 39; DB 3; Length 2105;
Best Local Similarity 53.8%; Pred. No. 5.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YOHDLRAYGFWRLL 15
|||:|:|:|:
Db 383 YOHVLOAAGPWHM 395

RESULT 24
US-09-428-371-3
; Sequence 3, Application US/09428371
; Patent No. 6800435

;; GENERAL INFORMATION:
;; APPLICANT: Soderlund, David M.
;; APPLICANT: Knipple, Douglas C.
;; APPLICANT: Ingles, Patricia J.

;; TITLE OF INVENTION: INSECT SODIUM CHANNELS FROM INSECTICIDE-SUSCEPTIBLE AND
;; TITLE OF INVENTION: INSECTICIDE-RESISTANT HOUSE FLIES
;; CURRENT APPLICATION NUMBER: US/09/428,371
;; CURRENT FILING DATE: 1999-10-28

;; EARLIER APPLICATION NUMBER: 08/608,618
;; EARLIER FILING DATE: 1996-03-01
;; EARLIER APPLICATION NUMBER: 08/772,512
;; EARLIER FILING DATE: 1996-12-24
;; NUMBER OF SEQ ID NOS: 19
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 3

;; LENGTH: 2105
;; TYPE: PRT
;; ORGANISM: Musca domestica
US-09-428-371-3

Query Match 44.8%; Score 39; DB 4; Length 2105;
Best Local Similarity 53.8%; Pred. No. 5.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YOHDLRAYGFWRLL 15
|||:|:|:|:
Db 383 YOHVLOAAGPWHM 395

RESULT 25
US-08-936-165A-376
; Sequence 376, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:

APPLICANT: Black, Michael
APPLICANT: Burnham, Martin
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael
APPLICANT: Nicholas, Richard
APPLICANT: Pratt, Julie
APPLICANT: Reichard, Richard
APPLICANT: Rosenberg, Martin
APPLICANT: Ward, Judith
TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
TITLE OF INVENTION: Polypeptides and Their Uses
NUMBER OF SEQUENCES: 534
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936.165A
FILING DATE: 24-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/027,032
FILING DATE: 24-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmi, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50549
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 376:
SEQUENCE CHARACTERISTICS:
LENGTH: 66 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-936-165A-376
Query Match 43.7%; Score 38; DB 3; Length 66;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Qy 2 SYQHDLRAYGFWRL 15
Db 25 SYEPIEACGTWRL 38
RESULT 26
US-09-328-352-7204
Sequence 7204, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breston et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTCC9-03PA
CURRENT APPLICATION NUMBER: US/09/328.352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 7204
LENGTH: 109
TYPE: PRI

ORGANISM: Acinetobacter baumannii
US-09-328-352-7204
Query Match 43.7%; Score 38; DB 4; Length 109;
Best Local Similarity 42.9%; Pred. No. 32;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
Qy 1 RSYQHDLRAYGFWR 14
Db 41 KAYSQDLQRQGWK 54
RESULT 27
US-08-946-329A-65
Sequence 65, Application US/08946329A
Patent No. 6057091
GENERAL INFORMATION:
APPLICANT: Beachy, Philip A.
TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
NUMBER OF SEQUENCES: 109
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/946.329A
FILING DATE: 07-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/061,323
FILING DATE: 07-OCT-1996
APPLICATION NUMBER: 08/729,743
FILING DATE: 10-JUL-1996
APPLICATION NUMBER: 08/567,357
FILING DATE: 04-DEC-1995
APPLICATION NUMBER: 08/349,498
FILING DATE: 02-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/140001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 148 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-946-329A-65
Query Match 43.7%; Score 38; DB 3; Length 148;
Best Local Similarity 53.8%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 2 SYQHDLRAYGFWR 14
Db 105 SQDHDLEGRGWK 117
RESULT 28
US-09-107-532A-4183
Sequence 4183, Application US/09107532A
Patent No. 6583275

```

GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Denekle
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4183:
SEQUENCE CHARACTERISTICS:
LENGTH: 181 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...181
SEQUENCE DESCRIPTION: SEQ ID NO: 4183:
US-09-107-532A-4183

```

```

1  COUNTRY: USA
2  ZIP: 90067
3  COMPUTER READABLE FORM:
4  MEDIUM TYPE: Floppy disk
5  COMPUTER: IBM PC compatible
6  OPERATING SYSTEM: PC-DOS/MS-DOS
7  SOFTWARE: PatentIn Release #1.0, Version #1.25
8  CURRENT APPLICATION DATA:
9  APPLICATION NUMBER: US/08/362,739
10 FILING DATE:
11 CLASSIFICATION: 435
12 PRIOR APPLICATION DATA:
13 APPLICATION NUMBER: US 08/040,747
14 FILING DATE: 31-MAR-1993
15 ATTORNEY/AGENT INFORMATION:
16 NAME: Wetherell, Jr. Ph.D., John R.,
17 REGISTRATION NUMBER: 31,678
18 REFERENCE/DOCKET NUMBER: PD-2097
19 TELECOMMUNICATION INFORMATION:
20 TELEPHONE: (619) 455-5100
21 TELEFAX: (619) 455-5110
22 INFORMATION FOR SEQ ID NO: 2:
23 SEQUENCE CHARACTERISTICS:
24 LENGTH: 320 amino acids
25 TYPE: amino acid
26 TOPOLOGY: linear
27 MOLECULE TYPE: protein
28 US-08-362-739-2
29
30 Query Match 43.7%; Score 38; DB 1; Length 320;
31 Best Local Similarity 46.2%; Pred. No. 1e+02;
32 Matches 6; Conservative 2; Mismatches 5; Indels
33
34 QY 3 YQDLRAYGFWRLL 15
35 | : | | | | :
36 Db 131 YESDFGKYFFWRV 143
37
38 RESULT 30
39 US-08-914-350A-2
40 Sequence 2, Application US/08914350A
41 Patent No. 6685945
42 GENERAL INFORMATION:
43 APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
44 APPLICANT: HAARE, David A.
45 APPLICANT: BLANCO, David R.
46 APPLICANT: CHAMPION, Cheryl I.
47 APPLICANT: LOVETT, Michael A.
48 APPLICANT: MILLER, James N.
49 TITLE OF INVENTION: CLONED LEPTOSPIRA OUTER MEMBRANE PROTEIN
50 FILE REFERENCE: UCLAI420-2
51 CURRENT APPLICATION NUMBER: US/08/914,350A
52 CURRENT FILING DATE: 1997-08-19
53 PRIOR APPLICATION NUMBER: US 08/362,739
54 PRIOR FILING DATE: 1994-12-20
55 PRIOR APPLICATION NUMBER: US 08/040,747
56 PRIOR FILING DATE: 1993-03-31
57 NUMBER OF SEQ ID NOS: 7
58 SOFTWARE: PatentIn version 3.1
59 SEQ ID NO 2
60 LENGTH: 320
61 TYPE: PRT
62 ORGANISM: Leptospira kirschneri
63 US-08-914-350A-2

```

APPLICANT: Miller, James N.
TITLE OF INVENTION: Cloned Leptospira Outer Membrane Protein
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Jubas & Lubitz
STREET: 1880 Century Park East, Suite 500
CITY: Los Angeles
STATE: California

FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 13652
LENGTH: 827
TYPE: PRT
ORGANISM: Myxococcus xanthus
US-09-902-540-13652

Query Match 43.7%; Score 38; DB 4; Length 827;
Best Local Similarity 60.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 QHDLRAYFW 13
Db 644 ERKLAYGLW 653

RESULT 34
US-08-317-450B-15
Sequence 15, Application US/08317450B
Patent No. 5660982
GENERAL INFORMATION:
APPLICANT: Trygsvason, Karl
APPLICANT: Kallunki, Pekka
APPLICANT: Pyke, Charles
TITLE OF INVENTION: Laminin Chains: Diagnostic and
TITLE OF INVENTION: Therapeutic Use
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: BANNER & ALLEGRETTI, LTD.
STREET: Ten South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/317,450B
FILING DATE: 04-OCT-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Chao, Mark
REGISTRATION NUMBER: 37,293
REFERENCE/DOCKET NUMBER: 94,778
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1111 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-317-450B-15

Query Match 43.7%; Score 38; DB 1; Length 1111;
Best Local Similarity 87.5%; Pred. No. 3.9e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RSYQHDLR 8
Db 843 RSYQHSR 850

RESULT 35

US-09-902-540-16548
Sequence 16548, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 16548
LENGTH: 557
TYPE: PRT
ORGANISM: Myxococcus xanthus
US-09-902-540-16548

Query Match 43.7%; Score 38; DB 4; Length 557;
Best Local Similarity 77.8%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 DLAYGFWR 14
Db 334 DLLARGFWR 342

RESULT 32
US-09-252-991A-28831
Sequence 28831, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 28831
LENGTH: 681
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28831

Query Match 43.7%; Score 38; DB 4; Length 681;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RSYQHDRLAYGF 12
Db 440 REFAEDIRMYGF 451

RESULT 33
US-09-902-540-13652
Sequence 13652, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

US-08-800-593-15
; Sequence 15, Application US/08800593
; Patent No. 6143505
; GENERAL INFORMATION:
; APPLICANT: Tryggvason, Karl
; APPLICANT: Kallunki, Pekka
; APPLICANT: Pyke, Charles
; TITLE OF INVENTION: Laminin Chains: Diagnostic and
; TITLE OF INVENTION: Therapeutic Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/800,593
; FILING DATE: 18-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/317,450
; FILING DATE: 04-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Chao, Mark
; REGISTRATION NUMBER: 37,293
; REFERENCE/DOCKET NUMBER: 94,778-B
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1111 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-800-593-15

Query Match 43.7%; Score 38; DB 3; Length 1111;
Best Local Similarity 87.5%; Pred. No. 3.9e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RSYQHDLR 8
Db 843 RSYQHSLR 850

RESULT 36
US-08-445-135-1
; Sequence 1, Application US/08445135
; Patent No. 5658789
; GENERAL INFORMATION:
; APPLICANT: Quaranta, Vito
; APPLICANT: Hormia, Marketta
; TITLE OF INVENTION: Promotion of Epithelial Cell Adhesion
; TITLE OF INVENTION: and Hemidesmosome Assembly by a Laminin-Like Molecule
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobber, Martens, Olson and Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,135
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/151,134
; FILING DATE: 12-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: DESMOS.002CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1171 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; US-08-445-135-1

Query Match 43.7%; Score 38; DB 1; Length 1171;
Best Local Similarity 87.5%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RSYQHDLR 8
Db 822 RSYQHSLR 829

RESULT 37
US-09-560-385A-36
; Sequence 36, Application US/09560385A
; Patent No. 6703363
; GENERAL INFORMATION:
; APPLICANT: Boutand, Ariel
; TITLE OF INVENTION: Recombinant Laminin 5
; FILE REFERENCE: 99-274-C
; CURRENT APPLICATION NUMBER: US/09/560,385A
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 36
; LENGTH: 1171
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-560-385A-36

Query Match 43.7%; Score 38; DB 4; Length 1171;
Best Local Similarity 87.5%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RSYQHDLR 8
Db 823 RSYQHSLR 830

RESULT 38
US-09-560-385A-28
; Sequence 28, Application US/09560385A
; Patent No. 6703363
; GENERAL INFORMATION:
; APPLICANT: Boutand, Ariel
; TITLE OF INVENTION: Recombinant Laminin 5
; FILE REFERENCE: 99-274-C

Search completed: March 18, 2005, 18:32:20
Job time : 43 secs

; CURRENT APPLICATION NUMBER: US/09/560,385A
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 1172
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-560-385A-28

Query Match 43.7%; Score 38; DB 4; Length 1172;
Best Local Similarity 87.5%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RSYQHDLR 8
|||||
Db 822 RSYQHSLR 829

RESULT 39
US-09-560-385A-32
; Sequence 32, Application US/09560385A
; Patent No. 6703363
; GENERAL INFORMATION:
; APPLICANT: Boutand, Ariel
; TITLE OF INVENTION: Recombinant Laminin 5
; FILE REFERENCE: 99-274-C
; CURRENT APPLICATION NUMBER: US/09/560,385A
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 1172
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-560-385A-32

Query Match 43.7%; Score 38; DB 4; Length 1172;
Best Local Similarity 87.5%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RSYQHDLR 8
|||||
Db 822 RSYQHSLR 829

RESULT 40
US-09-560-385A-34
; Sequence 34, Application US/09560385A
; Patent No. 6703363
; GENERAL INFORMATION:
; APPLICANT: Boutand, Ariel
; TITLE OF INVENTION: Recombinant Laminin 5
; FILE REFERENCE: 99-274-C
; CURRENT APPLICATION NUMBER: US/09/560,385A
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 1192
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-560-385A-34

Query Match 43.7%; Score 38; DB 4; Length 1192;
Best Local Similarity 87.5%; Pred. No. 4.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RSYQHDLR 8
|||||
Db 844 RSYQHSLR 851

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 18:14:26 ; Search time 38 Seconds
(without alignments)
37.980 Million cell updates/sec

Title: US-09-623-038-6

Perfect score: 87

Sequence: 1 RSYQHDLRAYGFWRL 15

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 43 | 49.4 | 118 | 2 | G81992 |
| 2 | 43 | 49.4 | 118 | 2 | G81221 |
| 3 | 43 | 49.4 | 188 | 2 | T02822 |
| 4 | 43 | 49.4 | 459 | 2 | D81950 |
| 5 | 42 | 48.3 | 444 | 2 | T24844 |
| 6 | 42 | 48.3 | 887 | 2 | T03939 |
| 7 | 42 | 48.3 | 994 | 1 | SYNCLM |
| 8 | 41 | 47.1 | 188 | 2 | D82248 |
| 9 | 41 | 47.1 | 817 | 2 | T43659 |
| 10 | 40 | 46.0 | 223 | 2 | A84436 |
| 11 | 40 | 46.0 | 274 | 2 | F86961 |
| 12 | 40 | 46.0 | 330 | 2 | B96027 |
| 13 | 40 | 46.0 | 371 | 2 | T21153 |
| 14 | 40 | 46.0 | 414 | 2 | T19912 |
| 15 | 40 | 46.0 | 447 | 2 | E72320 |
| 16 | 40 | 46.0 | 451 | 2 | T15718 |
| 17 | 40 | 46.0 | 516 | 2 | D36682 |
| 18 | 40 | 46.0 | 912 | 2 | T31223 |
| 19 | 40 | 46.0 | 1177 | 2 | AG0284 |
| 20 | 39.5 | 45.4 | 309 | 2 | A95988 |
| 21 | 39.5 | 45.4 | 471 | 2 | D95940 |
| 22 | 39 | 44.8 | 81 | 2 | A75372 |
| 23 | 39 | 44.8 | 103 | 2 | A53461 |
| 24 | 39 | 44.8 | 206 | 2 | S40829 |
| 25 | 39 | 44.8 | 206 | 2 | G86076 |
| 26 | 39 | 44.8 | 206 | 2 | H91229 |
| 27 | 39 | 44.8 | 286 | 2 | A55330 |
| 28 | 39 | 44.8 | 313 | 2 | AF1973 |
| 29 | 39 | 44.8 | 330 | 2 | AB2586 |

30 39 44.8 330 2 B97368 hypothetical sugar
31 39 44.8 335 2 I36790 hypothetical prote
32 39 44.8 340 2 C83827 molybdopterin bios
33 39 44.8 433 1 B65058 fixC protein homol
34 39 44.8 502 2 B75287 sensor histidine k
35 39 44.8 505 2 F71266 probable tpr prote
36 39 44.8 514 2 H87465 sodium-galactoside
37 39 44.8 700 2 S09699 bib protein - frui
38 39 44.8 736 2 T06757 hypothetical prote
39 39 44.8 978 2 H81311 transcription-repa
40 39 44.8 1055 2 T31111 ATPase 1 (EC 3.6.1
41 39 44.8 1135 1 J01928 G2-G1 polypeptid
42 39 44.8 2108 2 S72458 sodium channel pro
43 38.5 44.3 159 2 AH2118 bacterioferritin c
44 38.5 44.3 307 2 F97639 hypothetical prote
45 38.5 44.3 307 2 AG2862 conserved hypothet

ALIGNMENTS

RESULT 1

G81992

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) I chain A NMA0019 [similarity] - Neisseria

C:Species: Neisseria meningitidis

C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004

C:Accession: G81992

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel

; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000

A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.

A:Reference number: A81775; MUID:2022556; PMID:10761919

A:Accession: G81992

A>Status: preliminary

A:Residues: 1-118 <PAR>

A:Cross-references: UNIPROT:Q9JX78; GB:AL162752; GB:AL157959; NID:g7378778; PIDN:CA88333

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: nuoA; NMA0019

A:Superfamily: NADH dehydrogenase (ubiquinone) chain 3

C:Keywords: membrane-associated complex; NAD; oxidoreductase

Query Match 49.4%; Score 43; DB 2; Length 118;
Best Local Similarity 70.0%; Pred. No. 5;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 DLRAYGFWRL 15

||| ||||| :

Db 86 DLGAYGFWSM 95

RESULT 2

G81221

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) I chain A NMB0241 [similarity] - Neisseria

C:Species: Neisseria meningitidis

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004

C:Accession: G81221

R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.,

ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;

A>Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755; PMID:10710307

A:Accession: G81221

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-118 <TET>

A:Cross-references: UNIPROT:Q9K1C3; GB:AE002381; GB:AE002098; NID:g7225455; PIDN:AAF4065

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB0241

C:Superfamily: NADH dehydrogenase (ubiquinone) chain 3
C:Keywords: membrane-associated complex; NAD; oxidoreductase

Query Match 49.4%; Score 43; DB 2; Length 118;
Best Local Similarity 70.0%; Pred. No. 5;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 DLRAYGFWRLL 15
||| ||||| :
Db 86 DLGAYGFWSM 95

RESULT 3

T02822
probable membrane protein L2759.14 [imported] - Leishmania major (strain Friedlin)
C:Species: Leishmania major
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C:Accession: B01459; T02822
R:Myler, P.J.; Audleman, L.; deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magnus, C.;
Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
A:Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c
A:Reference number: B01455; MUID:99178987; PMID:10077609
A:Accession: B01459
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-188 <PYL>
A:Cross-references: UNIPROT:Q25352; GB:AE001274; NID:g3264850; PIDN:AAC24644.1; PID:g140
A:Experimental source: strain MHOM/IL/81/Friedlin
C:Genetics:
A:Gene: L2759.14
A:Map position: 1
C:Keywords: transmembrane protein

Query Match 49.4%; Score 43; DB 2; Length 188;
Best Local Similarity 77.8%; Pred. No. 8.1;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 5 HDLRAYGF 13
||| |||||
Db 46 HDLMAIGFW 54

RESULT 4

D81950
probable transmembrane efflux protein NMA1022 [imported] - Neisseria meningitidis (strain
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: D81950
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: D81950
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-459 <PAR>
A:Cross-references: UNIPROT:Q9JV27; GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CAB8429
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA1022
C:Superfamily: conserved hypothetical protein H11612

Query Match 49.4%; Score 43; DB 2; Length 459;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 SYQHDRLRAYGF 13
||| : |||||
Db 416 AYRDMGIYGF 427

RESULT 5

T24844

hypothetical protein T11F9.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T24844
R:Lennard, N.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19941
A:Accession: T24844
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-444 <WIL>
A:Cross-references: UNIPROT:Q22405; EMBL:Z74042; PIDN:CAA98536.1; GSPDB:GN00023; CBSP:T.1
A:Experimental source: clone T11F9
C:Genetics:
A:Gene: CESP:T11F9.1
A:Map position: 5
A:Introns: 19/1; 61/2; 130/3; 177/3; 347/3; 375/1; 413/3

Query Match 48.3%; Score 42; DB 2; Length 444;
Best Local Similarity 42.9%; Pred. No. 29;
Matches 9; Conservative 3; Mismatches 3; Indels 6; Gaps 1;

Qy 1 RSYQHDLEA-----YGFWRLL 15
||| : ||| : ||| :
Db 189 REYQLNLSALARYEYGVWRV 209

RESULT 6

T03939
potassium channel protein - maize
C:Species: Zea mays (maize)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T03939
R:Hoch, S.; Dreyer, I.; Dietrich, P.; Becker, D.; Mueller-Roeber, B.; Hedrich, R.
Proc. Natl. Acad. Sci. U.S.A. 94, 4806-4810, 1997
A:Title: Molecular basis of plant-specific acid activation of K⁺ uptake channels.
A:Reference number: Z14922; MUID:97272307; PMID:9114073
A:Accession: T03939
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-887 <HOT>

A:Cross-references: UNIPROT:O04242; EMBL:Y07632; NID:g2104907; PIDN:CAA68912.1; PID:g210
A:Experimental source: strain L551311B, sub species Apache, coleoptile
C:Superfamily: Arabidopsis potassium channel protein AKT1; ankyrin repeat homology; CAMP
C:Keywords: potassium channel; transmembrane protein; transport protein; voltage-gated 1
F:561-593/Domain: ankyrin repeat homology <AN1>
F:658-690/Domain: ankyrin repeat homology <AN2>

Query Match 48.3%; Score 42; DB 2; Length 887;
Best Local Similarity 56.2%; Pred. No. 60;
Matches 9; Conservative 1; Mismatches 0; Indels 6; Gaps 1;

Qy 6 DLRAYGF-----WRL 15
||| : ||| : ||| :
Db 163 DLRSYGFNMLRLWRL 178

RESULT 7

SYNCLIM
leucine-tRNA ligase (EC 6.1.1.4) precursor, mitochondrial - Neurospora crassa
N:Alternate names: leucyl-tRNA synthetase
C:Species: Neurospora crassa
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C:Accession: A33474
R:Chow, C.M.; Metzberg, R.L.; Rajbhandary, U.L.
Mol. Cell. Biol. 9, 4631-4644, 1989
A:Title: Nuclear gene for mitochondrial leucyl-tRNA synthetase of Neurospora crassa: iso
A:Reference number: A33474; MUID:90097874; PMID:2574823
A:Accession: A33474
A:Molecule type: DNA
A:Residues: 1-994 <CHO>

A:Cross-references: UNIPROT:P15181; GB:M30472; NID:g168829; PIDN:AAA33599.1; PID:g168830

C:Genetics:
A:Gene: leu-5
C:Superfamily: leucine-tRNA ligase
C:Keywords: aminoacyl-tRNA synthetase; ATP; ligase; mitochondrion; protein biosynthesis;
F:1-52/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F:53-994/Product: leucine-tRNA ligase #status predicted <MAT>
F:100-103/Region: ATP binding #status predicted
F:503-550/Region: zinc finger CCCC motif

Query Match 48.3%; Score 42; DB 1; Length 994;
Best Local Similarity 60.0%; Pred. No. 68;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 HDLRAYGFWR 14
||| : ||:
DB 403 HDLRDHAFWK 412
||| : ||:

RESULT 8
DB2248
RNA polymerase sigma-70 factor, ECP subfamily VC1045 [imported] - Vibrio cholerae (strain
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: DB2248
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardon, D.; Emolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, P.
I, R.K.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: AB2035; MUID:20406833; PMID:10952301
A:Accession: DB2248
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-188 <HEI>
A:Cross-references: UNIPROT:Q9KT60; GB:AE004186; GB:AE003852; NID:g9655507; PIDN:AAF9420
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1045
A:Map position: 1

Query Match 47.1%; Score 41; DB 2; Length 188;
Best Local Similarity 46.2%; Pred. No. 17;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 RSYQHDRLRAYGFW 13
||| : ||:
DB 32 RAYHRDLVRYAYW 44
||| : ||:

RESULT 9
T43659
trehalose-phosphatase (EC 3.1.3.12) [similarity] - fission yeast (Schizosaccharomyces po
N:Alternate names: trehalose-6P phosphatase
C:Species: Schizosaccharomyces pombe
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T43659; T38002; T43056
R:Canasado, J.; Vicente, J.; Soto, T.; Franco, A.; Castillo, R.; Gacto, M.
submitted to the EMBL Data Library, June 1999
A:Description: Isolation and characterization of trehalose-6P phosphatase disrupted muta
A:Reference number: Z22608.
A:Accession: T43659
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-817 <CAN>
A:Cross-references: UNIPROT:P78875; EMBL:AJ242743; PIDN:CAB45142.1
A:Experimental source: strain 972h(-)
R:Olliver, K.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1997
A:Reference number: Z21760
A:Accession: T38002
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-817 <OLI>

A:Cross-references: EMBL:Z97209; PIDN:CAB10126.1; GSPDB:GN00066; SPDB:SPAC19G12.15C
A:Experimental source: strain 972h-; cosmid cl9G12
R:Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
DNA Res. 4, 363-369, 1997
A:Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
A:Reference number: Z17323; MUID:98162722; PMID:9501991
A:Accession: T43056
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 505-817 <YOS>
A:Cross-references: EMBL:DB9225; NID:gi749657; PIDN:BAAL3886.1; PID:gi749658
A:Experimental source: strain PR745
C:Genetics:
A:Gene: cppl; SPAC19G12.15C
A:Map position: 1
C:Function:
A:Description: catalyzes the hydrolysis of trehalose 6-phosphate to trehalose
A:Pathway: trehalose biosynthesis
C:Keywords: phosphoric monoester hydrolase; trehalose biosynthesis

Query Match 47.1%; Score 41; DB 2; Length 817;
Best Local Similarity 57.1%; Pred. No. 81;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 RSYQHDRLRAYGFWR 14
||| : ||:
DB 362 RSFQRLRTYPEWR 375
||| : ||:

RESULT 10
AB4436
probable glutathione S-transferase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: AB4436
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, B.
euss, D.; Niernman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: AB4420; MUID:20083487; PMID:10617197
A:Accession: AB4436
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-223 <STO>
A:Cross-references: UNIPROT:Q9ZVQ4; GB:AE002093; NID:g3894170; PIDN:AAC78520.1; GSPDB:G1
C:Genetics:
A:Gene: At2G02380
A:Map position: 2
C:Superfamily: plaiice glutathione transferase

Query Match 46.0%; Score 40; DB 2; Length 223;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YQHDRLRAYGFWR 14
||| : ||:
DB 8 YQAKLKLVSYYR 19
||| : ||:

RESULT 11
F96961
bacitracin resistance protein (baca) [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C:Accession: F96961
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl.
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: F96961
A:Status: preliminary

```

A; Introns: 173/2; 211/3; 247/2

Query Match          46.0%; Score 40; DB 2; Length 371;
Best Local Similarity 87.5%; Pred. No. 52;
Matches              7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      8 RAYGFWRL 15
      ||||| |
DB      70 RAYGFWIL 77

RESULT 14
T19912
hypothetical protein C43F9.4 - Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C; Accession: T19912
R; Mortimore, B.
submitted to the EMBL Data Library, November 1996
A; Reference number: Z19195
A; Accession: T19912
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Residues: 1-414 <WIL>
A; Cross-references: UNIPROT:Q9XUG7; EMBL:Z82262; PIDN: CAB05151.1; GSPDB: GNO00022
C; Genetics:
A; Gene: CRSP:C43F9.4
A; Map position: 4
A; Introns: 53/3; 81/2; 111/3; 134/3; 164/3; 219/1; 346/1

Query Match          46.0%; Score 40; DB 2; Length 414;
Best Local Similarity 60.0%; Pred. No. 59;
Matches              6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      4 QHDLRAYGEW 13
      : | | | |
DB      21 KHSIMFYGEW 30

RESULT 15
E72320
hypothetical protein TW0884 - Thermotoga maritima (strain MSB8)
C; Species: Thermotoga maritima
C; Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C; Accession: E72320
R; Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Garret, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, C.M.
Nature 399, 323-329, 1999
A; Title: Evidence for lateral gene transfer between Archaea and Bacteria from 9
A; Reference number: A72200; MUID: 99287316; PMID: 10360571
A; Accession: E72320
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-447 <ARN>
A; Cross-references: UNIPROT:Q9WZV6; GB: AE001754; GB: AE000512; NID: g4981417; PID:
A; Experimental source: strain MSB8
C; Genetics:
A; Gene: TW0884

Query Match          46.0%; Score 40; DB 2; Length 447;
Best Local Similarity 66.7%; Pred. No. 63;
Matches              6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      5 HDLRAYGEW 13
      || | | |
DB      11 HDFAAYDEFW 19

RESULT 16
T15718
hypothetical protein C30G12.1 - Caenorhabditis elegans

```

C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15718
R:Latreille, P.
A:Description: The sequence of C. elegans cosmid C30G12.
A:Reference number: Z18393
A:Accession: T15718
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-451 <LAT>
A:Cross-references: EMBL:U21319; NID:G687832; PID:G687833; PIDN:AAC46670.1; CESP:C30G12.
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:C30G12.1
A:introns: 37/3; 71/2; 107/3; 153/2; 205/1; 289/1; 376/3

Query Match 46.0%; Score 40; DB 2; Length 451;
Best Local Similarity 77.8%; Pred. No. 64;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 SYQHDLRAY 10
DB 61 SYQHDLSAF 69
|||||

RESULT 17
D96682
C:Species: Arabidopsis thaliana
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: D96682
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Huiziar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maif, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96682
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-516 <STO>
A:Cross-references: UNIPROT:Q9SHX4; GB:AE005173; NID:G6686400; PIDN:AAF23834.1; GSPDB:GN
C:Genetics:
A:Gene: F1E22.18
A:Map position: 1

Query Match 46.0%; Score 40; DB 2; Length 516;
Best Local Similarity 54.5%; Pred. No. 74;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 YQHDLRAYGF 13
DB 115 YQHDLESYGLF 125
|||||

RESULT 18
T31223
C:Species: Sphingomonas aromaticivorans
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T31223
R:Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.; submitted to the EMBL Data Library, July 1998
A:Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas aromaticivorans
A:Reference number: Z20992
A:Accession: T31223
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-912 <ROM>
A:Cross-references: UNIPROT:O85931; EMBL:AF079317; NID:G3378261; PID:G3378364; PIDN:AAD
C:Genetics:
A:Genome: plasmid pNL1
A:Note: tag

Query Match 46.0%; Score 40; DB 2; Length 912;
Best Local Similarity 53.3%; Pred. No. 1.3e+02;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 RSYQHDLRAYGF 15
DB 537 RGYATDLRSQGWYL 551
|||||

RESULT 19
AG0284
A:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AG0284
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AG0284
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1177 <KUR>
A:Cross-references: UNIPROT:Q82B50; GB:AL590842; PIDN:CAC91139.1; PID:G15980331; GSPDB:K
C:Genetics:
A:Gene: nifJ
C:Superfamily: pyruvate (flavodoxin) dehydrogenase; ferredoxin 2[4Fe-4S] homology

Query Match 46.0%; Score 40; DB 2; Length 1177;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 9; Conservative 1; Mismatches 4; Indels 4; Gaps 1;

QY 2 SYQHD---LRAYGF 15
DB 1082 AYSHDQMKQLTATGF 1099
|||||

RESULT 20
A95988
A:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: A95988
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Herna: Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo: A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: A95988
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-309 <KUR>
A:Cross-references: UNIPROT:Q92UG2; GB:AL591985; PIDN:CAC49569.1; PID:G15141056; GSPDB:K
R:Galibert, F.; Finan, T.M.; Long, S.R.; Fuhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, pla, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T. Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yen, K.; A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMB20875

A:Genome: plasmid

Query Match 45.4%; Score 39.5; DB 2; Length 309;

Best Local Similarity 61.5%; Pred. No. 52; Mismatches 2; Indels 1; Gaps 1;

QY 3 YOHLDRAYGFWR 15

DB 76 YEYGARA-GFWRL 87

RESULT 21

D95940
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004

A:Accession: D95940

R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan

A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo

A:Reference number: A95842; MUID:21396508; PMID:11481431

A:Accession: D95940

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-471 <KUR>

A:Cross-references: UNIPROT:Q92VC0; GB:AL591985; PIDN:CAC49188.1; PID:g15140673; GSPDB:G

A:Experimental source: strain 1021, megaplasmid pSymB

R:Galibert, P.; Finan, T.M.; Long, S.R.; Uehler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: SMB21284

A:Genome: plasmid

C:Keywords: oxidoreductase

Query Match 45.4%; Score 39.5; DB 2; Length 471;

Best Local Similarity 61.5%; Pred. No. 81; Mismatches 2; Indels 1; Gaps 1;

QY 3 YOHLDRAYGFWR 15

DB 72 YEYGARA-GFWRL 83

RESULT 22

A75372

hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

A:Accession: A75372

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: A75372

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-81 <WHI>

A:Cross-references: UNIPROT:Q9RTW0; GB:AE002007; GB:AE000513; NID:g6459402; PIDN:AAF1120

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR1643

A:Map position: 1

Query Match 44.8%; Score 39; DB 2; Length 81;

Best Local Similarity 100.0%; Pred. No. 16;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 AYGFWR 14

DB 64 AYGFWR 69

RESULT 23

A53461

voltage-sensitive sodium channel - house fly (fragment)

C:Species: Musca domestica (house fly)

C:Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

A:Accession: A53461

R:Knipple, D.C.; Doyle, K.E.; Marsella-Herrick, P.A.; Soderlund, D.M.

Proc. Natl. Acad. Sci. U.S.A. 91, 2483-2487, 1994

A:Title: Tight genetic linkage between the kdr insecticide resistance trait and a voltage

A:Reference number: A53461; MUID:94195766; PMID:8146143

A:Contents: NAIDM, insecticide-susceptible

A:Accession: A53461

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-103 <KNI>

A:Cross-references: UNIPROT:Q25439; UNIPROT:Q25440; UNIPROT:Q94615

A:Note: sequence inconsistent with nucleotide translation

A:Note: sequence extracted from NCBI backbone (NCBIN:146081, NCBI:146080)

C:Superfamily: sodium channel protein

C:Keywords: duplication

Query Match 44.8%; Score 39; DB 2; Length 103;

Best Local Similarity 53.8%; Pred. No. 20;

Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YOHLDRAYGFWR 15

DB 46 YOHLDRAYGFWR 58

RESULT 24

S40829

hypothetical 23.5K protein (glnA-fdhE intergenic region) - Escherichia coli (strain K-12)

N:Alternate names: hypothetical protein o206

C:Species: Escherichia coli

C:Date: 06-Oct-1994 #sequence_revision 26-May-1995 #text_change 09-Jul-2004

A:Accession: S40829; H65193

R:Plunkett III, G.; Burland, V.; Daniels, D.L.; Blattner, F.R.

Nucleic Acids Res. 21, 3391-3396, 1993

A:Title: Analysis of the Escherichia coli genome. III. DNA sequence of the region from 8

A:Reference number: S40802; MUID:93347969; PMID:8346018

A:Accession: S40829

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-206 <PLU>

A:Cross-references: UNIPROT:P32145; EMBL:L19201; NID:g304961; PIDN:AAB03018.1; PID:g3049

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1993

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: H65193

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-206 <BLAT>

A:Cross-references: GB:AE000464; GB:U00096; NID:g2367324; PIDN:AAD13447.1; PID:gl790318;

A:Experimental source: strain K-12, substrain MG1655

C:Genetics:

A:Gene: yihX

A:Start codon: GTG

C:Superfamily: Alcaligenes eutrophus phosphoglycolate phosphatase

Query Match 44.8%; Score 39; DB 2; Length 206;

Best Local Similarity 66.7%; Pred. No. 42;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RSYQHDLRAYGF 12
| | | | |
Db 152 RIYQHVLAQEGF 163

RESULT 25
G86076
probable phosphatase yihX [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: G86076
R:Perna, N.R.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: G86076
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-206 <STO>
A:Cross-references: UNIPROT:Q8X8D1; GB:AE005174; NID:G12518771; PIDN:AGS9075.1; GSPDB:G
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yihX
C:Superfamily: Alcaligenes eutrophus phosphoglycolate phosphatase

Query Match 44.8%; Score 39; DB 2; Length 206;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RSYQHDLRAYGF 12
| | | | |
Db 152 RIYQHVLAQEGF 163

RESULT 26
H91229
probable phosphatase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 05099
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: H91229
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: H91229
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-206 <HAY>
A:Cross-references: UNIPROT:Q8X8D1; GB:BA000007; PIDN:BA838231.1; PID:G13364284; GSPDB:G
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: Ecs4808
C:Superfamily: Alcaligenes eutrophus phosphoglycolate phosphatase

Query Match 44.8%; Score 39; DB 2; Length 206;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RSYQHDLRAYGF 12
| | | | |
Db 152 RIYQHVLAQEGF 163

RESULT 27
A55530
megakaryocyte growth and development factor, long form - human
N:Alternate names: MPL ligand, long form
C:Species: Homo sapiens (man)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 07-May-1999

C:Accession: A55530
R:Chang, M.; McIninch, J.; Basu, R.; Shutter, J.; Hsu, R.; Perkins, C.; Mar, V.; Suggs, E.
J. Biol. Chem. 270, 511-514, 1995
A:Title: Cloning and characterization of the human megakaryocyte growth and development
A:Reference number: A55530; MUID:95122483; PMID:7822271
A:Accession: A55530
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-286 <CHA>
A:Cross-references: GB:U17071
C:Genetics:
A:Gene: MGDF
A:Map position: 3q26.3
C:Keywords: alternative splicing; cytokine

Query Match 44.8%; Score 39; DB 2; Length 286;
Best Local Similarity 58.3%; Pred. No. 59;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 SYQHDLRAYGF 13
| | | | |
Db 151 SFQHLRGKQFW 162

RESULT 28
AF1973
hypothetical protein alr1337 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AF1973
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AF1973
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-313 <KUR>
A:Cross-references: UNIPROT:Q8YX78; GB:BA000019; PIDN:BA873294.1; PID:G17130684; GSPDB:G
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr1337
C:Superfamily: ADP-ribosylglycohydrolase

Query Match 44.8%; Score 39; DB 2; Length 313;
Best Local Similarity 50.0%; Pred. No. 64;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RSYQHDLRAYGF 14
| | | | |
Db 69 RGYSLDAIAHSFWR 82

RESULT 29
AB2586
pfkB family carbohydrate kinase [imported] - Agrobacterium tumefaciens (strain C58, Dupont
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AB2586
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, Y.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan,
Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
acer, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AB2586
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-330 <KUR>

A;Cross-references: UNIPROT:Q8U5J9; GB:AE008688; PIDN:AA141104.1; PID:g17738396; GSPDB:C
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu0079
A;Map position: circular chromosome

Query Match 44.8%; Score 39; DB 2; Length 330;
Best Local Similarity 54.5%; Pred. No. 68;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 YQHDRLRAYGFW 13
:|||||:
Db 90 FQHDRAQGVY 100

RESULT 30

B97368
hypothetical sugar kinase slr0537 [imported] - Agrobacterium tumefaciens (strain C58, Ce
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: B97368
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: B97368
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-330 <NR>
A;Cross-references: UNIPROT:Q8U5P4; GB:AE007869; PIDN:AAK85899.1; PID:g15154944; GSPDB:C
C;Genetics:
A;Gene: AGR_C118
A;Map position: circular chromosome

Query Match 44.8%; Score 39; DB 2; Length 330;
Best Local Similarity 54.5%; Pred. No. 68;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 YQHDRLRAYGFW 13
:|||||:
Db 90 FQHDRAQGVY 100

RESULT 31

I36790
hypothetical protein ORF44 - ictaluriid herpesvirus 1 (strain auburn 1)
C;Species: ictaluriid herpesvirus 1
A;Note: host ictaluriid punctatus (channel catfish)
C;Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 09-Jul-2004
C;Accession: I36790
R;Davison, A.J.
A;Description: Channel catfish virus: a new type of herpesvirus.
A;Reference number: A36804
A;Accession: I36790
A;Molecule type: DNA
A;Residues: 1-335 <DAV>
A;Cross-references: UNIPROT:Q00124; GB:M75136; MUID:g331209; PIDN:AAA88147.1; PID:g331254
R;Davison, A.J.
Virology 186, 9-14, 1992
A;Title: Channel catfish virus: a new type of herpesvirus.
A;Reference number: A39447; MUID:92087490; PMID:1727613
A;Contents: annotation
A;Note: neither protein nor nucleic acid sequence is given
C;Genetics:
A;Gene: 44
C;Superfamily: ictaluriid herpesvirus 1 hypothetical protein ORF44

Query Match 44.8%; Score 39; DB 2; Length 335;
Best Local Similarity 63.6%; Pred. No. 69;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 HDLRAYGFWRL 15
:|||||:
Db 249 YDLRCGCGWRL 259

RESULT 32

C83827
molybdopterin biosynthesis moeB [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: C83827
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira,
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: C83827
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-340 <STO>
A;Cross-references: UNIPROT:Q9KD00; GB:AP001512; GB:BA000004; NID:g10174030; PIDN:BA0051
A;Experimental source: strain C-125
C;Genetics:
A;Gene: moeB

Query Match 44.8%; Score 39; DB 2; Length 340;
Best Local Similarity 35.7%; Pred. No. 70;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 RSYQHDRLRAYGFW 14
:|||||:
Db 215 QAYRHSLMSFDVWR 228

RESULT 33

B65058
fixC protein homolog b2766 - Escherichia coli (strain K-12)
N;Contains: probable quinone reductase (EC 1.6.5.-)
C;Species: Escherichia coli
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Mar-2002
C;Accession: B65058
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: B65058
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-433 <BLAT>
A;Cross-references: GB:AE000360; GB:U00096; NID:g2367157; PIDN:AA75808.1; PID:g1789125;
A;Experimental source: strain K-12, substrain MG1655
C;Superfamily: fixC protein
C;Keywords: oxidoreductase

Query Match 44.8%; Score 39; DB 1; Length 433;
Best Local Similarity 46.2%; Pred. No. 90;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 YQHDRLRAYGFWRL 15
:|||||:
Db 410 FWHHLRRHGLMHL 422

RESULT 34

B75287
sensor histidine kinase - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: B75287
R;White, O.; Eisen, J.A.; Heidelberg, J.P.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999

[illegible]

H81311
transcription-repair coupling factor Cj1085c [imported] - Campylobacter jejuni (strain N
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: H81311
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: H81311
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-978 <PAR>
A:Cross-references: UNIPROT:O9PNK9; GB:AL139077; GB:AL111168; NID:g6968444; PIDN:CAB7334
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: mfd; Cj1085c
C:Superfamily: transcription-repair coupling protein

Query Match 44.8%; Score 39; DB 2; Length 978;
Best Local Similarity 55.6%; Pred. No. 2.1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 HDLRAYGFW 13
||: : |||
Db 213 HDINSLGFW 221

RESULT 40
T31111
ATPase 1 (EC 3.6.1.-), P-type - yeast (Schwanniomyces occidentalis)
C:Species: Schwanniomyces occidentalis
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31111
R:Banuelos, M.A.; Rodriguez-Navarro, A.
J. Biol. Chem. 273, 1640-1646, 1998
A:Title: P-type ATPases mediate sodium and potassium effluxes in Schwanniomyces occiden
A:Reference number: Z11524; MUID:98104151; PMID:9430707
A:Accession: T31111
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1055 <BAN>
A:Cross-references: UNIPROT:O13397; EMBL:AF030860; NID:g2623235; PID:g2623236; PIDN:AAB8
C:Genetics:
A:Gene: ENA1
C:Function:
A:Description: mediates sodium and potassium effluxes
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding domain
C:Keywords: hydrolase
P:1616-796/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 44.8%; Score 39; DB 2; Length 1055;
Best Local Similarity 75.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 DLRAYGFW 13
|: |||||
Db 866 DMAYGFW 873

Search completed: March 18, 2005, 18:31:31
Job time : 38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 18:13:51 ; Search time 175 Seconds
(without alignments)
43.893 Million cell updates/sec

Title: US-09-623-038-6

Perfect score: 87

Sequence: 1 RSYQHDLRAYGFWR 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------|-------------|
| 1 | 47 | 54.0 | 533 | Q8RNY4 | hafnia alve |
| 2 | 46 | 52.9 | 227 | Q8T09 | methanosarc |
| 3 | 46 | 52.9 | 587 | Q86H19 | dictyosteli |
| 4 | 45 | 51.7 | 374 | Q68J85 | escherichia |
| 5 | 44 | 50.6 | 1056 | Q64N19 | bacteroides |
| 6 | 44 | 50.6 | 1430 | Q72Z88 | brachydanio |
| 7 | 44 | 50.6 | 1478 | THO2_HUMAN | homo sapien |
| 8 | 44 | 50.6 | 1504 | Q6DRE1 | brachydanio |
| 9 | 43 | 49.4 | 118 | Q9UX78 | neisseria m |
| 10 | 43 | 49.4 | 118 | Q9K1C3 | neisseria m |
| 11 | 43 | 49.4 | 188 | Q25352 | leishmania |
| 12 | 43 | 49.4 | 256 | Q82MF7 | streptomyce |
| 13 | 43 | 49.4 | 430 | Q6J925 | arthrolepti |
| 14 | 43 | 49.4 | 459 | 1_NORM_NEIMA | neisseria m |
| 15 | 43 | 49.4 | 465 | Q6UQ64 | bacillus st |
| 16 | 43 | 49.4 | 530 | Q69B20 | bacillus st |
| 17 | 42 | 48.3 | 318 | Q8TPC0 | methanosarc |
| 18 | 42 | 48.3 | 318 | Q742P0 | mycobactari |
| 19 | 42 | 48.3 | 321 | Q67P30 | mycobacte |
| 20 | 42 | 48.3 | 418 | Q8TRX2 | methanosarc |
| 21 | 42 | 48.3 | 444 | Q22405 | caenorhabdi |
| 22 | 42 | 48.3 | 652 | Q9VJH0 | drosophila |
| 23 | 42 | 48.3 | 880 | 1_SVA_BUCBP | buchnera ap |
| 24 | 42 | 48.3 | 887 | Q04242 | zea mays (m |
| 25 | 42 | 48.3 | 907 | Q8H569 | oryza sativ |
| 26 | 42 | 48.3 | 951 | Q7MUS9 | porphyromon |
| 27 | 42 | 48.3 | 994 | 1_SYL181_NEUCR | neurospora |
| 28 | 41.5 | 47.7 | 306 | Q9NEG4 | leishmania |
| 29 | 41 | 47.1 | 60 | Q721X3 | botryllus s |
| 30 | 41 | 47.1 | 61 | Q9H356 | homo sapien |
| 31 | 41 | 47.1 | 61 | Q8D683 | vibrio vuln |

| | | | | | | |
|----|----|------|-----|---|-----------|--------------------|
| 32 | 41 | 47.1 | 188 | 2 | Q9KT60 | Q9kt60 vibrio chol |
| 33 | 41 | 47.1 | 191 | 2 | Q7MIS3 | Q7mis3 vibrio vuln |
| 34 | 41 | 47.1 | 191 | 2 | Q8DB49 | Q8db49 vibrio vuln |
| 35 | 41 | 47.1 | 229 | 2 | Q8PW04 | Q8pwo4 methanosarc |
| 36 | 41 | 47.1 | 260 | 2 | Q8G977 | Q8g977 pseudomonas |
| 37 | 41 | 47.1 | 260 | 2 | Q8GBV1 | Q8gbv1 pseudomonas |
| 38 | 41 | 47.1 | 301 | 1 | RL5_NEUCR | RL5_NEUCR |
| 39 | 41 | 47.1 | 323 | 2 | Q6MAL4 | Q6mal4 parachlamyd |
| 40 | 41 | 47.1 | 338 | 2 | Q925V9 | Q925v9 pseudomonas |
| 41 | 41 | 47.1 | 354 | 2 | Q8ZVC3 | Q8zvc3 pyrobaculum |
| 42 | 41 | 47.1 | 368 | 2 | Q73TGI | Q73tgi mycobacteri |
| 43 | 41 | 47.1 | 404 | 2 | O52212 | O52212 pseudomonas |
| 44 | 41 | 47.1 | 404 | 2 | O6XUM9 | O6xum9 pseudomonas |
| 45 | 41 | 47.1 | 404 | 2 | Q7BV60 | Q7bv60 incl/m plas |

ALIGNMENTS

RESULT 1

Q8RNY4 PRELIMINARY; PRT; 533 AA.
AC Q8RNY4;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Restriction endonuclease.
GN Name=esp3IR;
OS Hafnia alvei.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Hafnia.
OX NCBI_TaxID=569;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RFU3;
RX MEDLINE=22162868; PubMed=12172806; DOI=10.1007/s00438-002-0701-6;
RA Bitinaite J., Mickaite G., Dauksaite V., Jakubauskas A., Timinskas A.,
RA Vaisvila R., Lubyas A., Janulaitis A.;
RT "Evolutionary relationship of Alw26I, Eco31I and Esp3I, restriction
RT endonucleases that recognise overlapping sequences.";
RL Mol. Genet. Genomics 267:664-672(2002).
DR EMBL; AF458983; AAM09641.1; -
DR GO; GO:0004519; F:endonuclease activity; IEA.
KW Endonuclease.
SQ SEQUENCE 533 AA; 61708 MW; 5691332C65A6AD60 CRC64;
Query Match 54.0%; Score 47; DB 2; Length 533;
Best Local Similarity 53.8%; Pred. No. 20;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

RESULT 2

Q8T09 PRELIMINARY; PRT; 227 AA.
AC Q8T09;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Hypothetical protein MA0634.
GN OrderedLocusNames=MA0634;
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238; DOI=10.1101/gr.223902;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,

RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
 RA Allen N., Naylor J., Stange-Thomann N., DeArellano K., Johnson R.,
 RA Linton L., McEwan P., McKernan K., Talaman J., Tirrell A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
 RA Hederich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.T.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.,
 RT "The Genome of Methanobrevibacterium acetivorans reveals extensive metabolic
 RT and physiological diversity.";
 RT Genome Res. 12:532-542(2002).
 DR EMBL; AE010725; AAM04076.1; -;
 DR Pfam; PF02592; DUF165; 1.
 DR TIGRFAMs; TIGR00697; DUF165; 1.
 KW Complete proteome.
 SQ SEQUENCE 227 AA; 25462 MW; 39E2EB615EB7CAB6 CRC64;
 Query Match 52.9%; Score 46; DB 2; Length 227;
 Best Local Similarity 61.5%; Pred. No. 12;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 Qy 2 SYQHDLRAYGFW 14
 Db 120 SYQHDLWAFNFW 132
 RESULT 3
 Q86H19 PRELIMINARY; PRT; 587 AA.
 ID Q86H19
 AC Q86H19
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26; Last annotation update)
 DE Similar to Dictyostelium discoideum (Slime mold). Phosphatidylinositol
 DE 3-kinase 3 (EC 2.7.1.137) (PI3-kinase) (PtdIns-3-kinase) (PI3K).
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
 RA Gleckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
 RA Lehmann B., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
 RA Tuggal B., Cox E., Quail M.A., Platzer M., Roenthal A., Noegel A.A.;
 RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
 RT Nature 419:79-85(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RA Baumgart C.;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC11072; AAC2597.2; -;
 DR GO; GO:0016301; P:kinase activity; IEA.
 DR GO; GO:0004497; P:monooxygenase activity; IEA.
 DR GO; GO:0016303; P:phosphatidylinositol 3-kinase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome P450.
 DR InterPro; IPR002401; EP4501.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
 KW Kinase.
 SQ SEQUENCE 587 AA; 66683 MW; 4B960B35F9379E06 CRC64;
 Query Match 52.9%; Score 46; DB 2; Length 587;
 Best Local Similarity 53.3%; Pred. No. 33;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 RSYQHDLRAYGFW 15
 Db 120 SYQHDLWAFNFW 132

Db 289 RSGRHTLRGCMYQL 303
 RESULT 4
 Q68J85 PRELIMINARY; PRT; 374 AA.
 ID Q68J85
 AC Q68J85
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE TrbB.
 GN Name=trbB;
 OS Escherichia coli.
 OG Plasmid pSERB1.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Dudley E.G., Abe C., Ghigo J.-M., Nataro J.P.;
 RT "Type IV pilus mediated adherence of the atypical enteroaggregate
 RT Escherichia coli strain CI096 to cultured cells and abiotic
 RT surfaces.";
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY686591; AAT94232.1; -;
 DR InterPro; IPR001854; Ribosomal L29.
 DR PROSITE; PS00579; RIBOSOMAL_L29; UNKNOWN_1.
 KW Plasmid.
 SQ SEQUENCE 374 AA; 41006 MW; ABAEB985A412B121 CRC64;
 Query Match 51.7%; Score 45; DB 2; Length 374;
 Best Local Similarity 63.6%; Pred. No. 31;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 Qy 4 QHDLRAYGFW 14
 Db 96 KHPKRYGFW 106
 RESULT 5
 Q64N19 PRELIMINARY; PRT; 1056 AA.
 ID Q64N19
 AC Q64N19
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Putative outer membrane protein probably involved in nutrient
 DE binding.
 GN ORFNames=BF4380;
 OS Bacteroides fragilis.
 OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
 OC Bacteroidaceae; Bacteroides.
 OX NCBI_TaxID=817;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YCH46;
 RA Kuwahara T., Yamashita A., Hirakawa H., Nakayama H., Toh H., Okada N.,
 RA Kuwahara S., Hattori M., Hayashi T., Ohnishi Y.;
 RT "Genomic analysis of Bacteroides fragilis reveals extensive DNA
 RT inversions regulating cell surface adaptation.";
 RT Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2004).
 DR EMBL; AP006841; BAD51118.1; -;
 SQ SEQUENCE 1056 AA; 117353 MW; F1B335D1D41B988A CRC64;
 Query Match 50.6%; Score 44; DB 2; Length 1056;
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 Qy 2 SYQHDLRAYGFW 13
 Db 388 SYENIRSGAFW 399

RESULT 6

Q72288 PRELIMINARY; PRT; 1430 AA.
 AC Q72288;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE SI:021018.2 (Novel protein similar to human THO complex 2 (THOC2))
 DE (Fragment).
 GN Name=thoc2; Synonyms=SI:b21018.2;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Babbage A.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBSJ databases.
 DR EMBL; AL627383; CAD60810.1; -;
 DR 2FIN; ZDB-GENE-030616-54; thoc2.
 FT NON TER 1
 SQ SEQUENCE 1430 AA; 163500 MW; A31ADC92D23216AB CRC64;
 Query Match 50.6%; Score 44; DB 2; Length 1430;
 Best Local Similarity 58.3%; Pred. No. 1.8e+02;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 YOHDLRAYGFWR 14

|||||
 434 YQHYRLYGQWK 435

RESULT 7

THO2_HUMAN
 ID THO2_HUMAN STANDARD; PRT; 1478 AA.
 AC Q8N127; Q8H816;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE THO complex subunit 2 (Tho2).
 GN Name=THOC2;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, AND INTERACTION WITH THE
 RP TREX COMPLEX.
 RX MEDLINE=22010388; PubMed=11979277; DOI=10.1038/417304a;
 RA Straesser K., Masuda S., Mason P., Pfannstiel J., Oppizzi M.,
 RA Rodriguez-Navarro S., Rondon A.G., Aguilar A., Struhl K., Reed R.,
 RA Hurt E.;
 RT "TREX is a conserved complex coupling transcription with messenger RNA
 RT export.";
 RL Nature 417:304-308(2002).
 RN [2]
 RP SEQUENCE OF 444-1162 FROM N.A. (ISOFORM 1).
 RC TISSUE=Placenta;
 EX PubMed=14702039; DOI=10.1038/ngl285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Ohyashi M., Nishi T., Shibahara K., Tanaka T., Ishii S.,
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
 RA Sugawara M., Takahashi N., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 RA Omura Y., Abe K., Kamiyama K., Katsuta N., Sato K., Tanikawa M.,
 RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y.,
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
 RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
 RA Nomura Y., Togiyama S., Komai F., Hara R., Takeuchi K., Arita M.,

Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Oktani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsuura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs";
 RL Nat. Genet. 36:40-45(2004).
 RN [3]
 RP SEQUENCE OF 1151-1478 FROM N.A. (ISOFORM 2).
 RC TISSUE=Salivary gland;
 RA Oetjenwaelder B., Obermaier B., Deutschenbaur S., Mewes H.-W., Weill B.,
 RA Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
 RL Submitted (AUG-2003) to the EMBL/GenBank/DBSJ databases.
 CC -!- FUNCTION: The THO/TREX complex is recruited to transcribed genes
 CC and travels with the RNA polymerase during elongation. It may
 CC physically link proteins that function in transcription and in RNA
 CC export.
 CC -!- SUBUNIT: Part of the heteromultimeric THO/TREX complex containing
 CC THOC1, THOC2, THOC3, THOC4 and NFX1/UA956.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q8N127-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q8N127-2; Sequence=VSP_008587, VSP_008588;
 CC Note=No experimental confirmation available;
 CC -!- SIMILARITY: Belongs to the THOC2 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF441770; AAM28436.1; -;
 DR EMBL; AK023659; BAB14630.1; ALT_INIT.
 DR EMBL; BX648654; CAE46196.1; -;
 DR Genew; HGNC:19073; THOC2.
 DR MIM; 300395; -;
 DR InterPro; IPR011001; Saposin_like.
 DR TIGRFAMs; TIGR01847; Bacteriocin_sig; 1.
 DR KW Alternative splicing; mRNA processing; mRNA splicing; mRNA transport;
 KW Nuclear protein; RNA-binding; Transport.
 FT DOMAIN 1149 1476 Lys-rich.
 FT DOMAIN 808 813 Nuclear localization signal (Potential).
 FT VARSPLIC 1311 1314 DSI -> VSTA (in isoform 2).
 FT VARSPLIC 1315 1478 Missing (in isoform 2).
 FT VARSPLIC 692 692 /FTID=VSP_008587.
 FT CONFLICT 1161 1161 F -> S (in Ref. 2).
 FT CONFLICT 1241 1241 E -> G (in Ref. 3).
 FT CONFLICT 1241 1241 E -> K (in Ref. 3).
 SQ SEQUENCE 1478 AA; 169580 MW; FE1CBCE28C560BCF CRC64;
 Query Match 50.6%; Score 44; DB 1; Length 1478;
 Best Local Similarity 58.3%; Pred. No. 1.9e+02;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 YOHDLRAYGFWR 14

```

Db GO; GO:0016491; F:oxidoreductase activity; IEA.
Dr GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
Dr InterPro; IPR000440; Oxidored q4.
Dr Pfam; PF00507; Oxidored_q4_1.
KW Complete proteome; NAD; Oxidoreductase; Quinone; Transmembrane.
SQ SEQUENCE 118 AA; 13584 MW; 38C401130EDDAE17 CRC64;

Query Match 49.4%; Score 43; DB 2; Length 118;
Best Local Similarity 70.0%; Pred. No. 20;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 DLRAYGFWR 15
Db 86 DLGAYGFWSM 95

RESULT 10
Q9K1C3 PRELIMINARY; PRT; 118 AA.
AC Q9K1C3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE NADH dehydrogenase I, A subunit.
GN OrderedLocusNames=NM00241;
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307; DOI=10.1126/science.287.5459.1809;
RA Tettelin H., Saunders N.J., Heidelberg J.F., Jeffries A.C.,
RA Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F.,
RA Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R.T., Peterson J.D.,
RA Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,
RA Dougherty B.A., Mason T.M., Ciecko A., Parksey D.S., Blair E.,
RA Cittone H., Clark E.B., Cotton M.D., Utterback T.R., Khouri H.M.,
RA Qin H., Vamathevan J.J., Gill J., Scarlato V., Maignani V., Pizza M.,
RA Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R.,
RA Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58.";
RL Science 287:1809-1815(2000).
CC -!- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
sulfur (Fe-S) centers, to quinones in the respiratory chain (By
similarity).
CC -!- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the complex I subunit 3 family.
DR EMBL; AE002381; AAP40695.1; -.
DR PIR; G81221; G81221.
DR TIGR; NM80241; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
Dr InterPro; IPR000440; Oxidored_q4.
Dr Pfam; PF00507; Oxidored_q4_1.
KW Complete proteome; NAD; Oxidoreductase; Quinone; Transmembrane.
SQ SEQUENCE 118 AA; 13598 MW; 21F58B3BAC579F84 CRC64;

Query Match 49.4%; Score 43; DB 2; Length 118;
Best Local Similarity 70.0%; Pred. No. 20;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 DLRAYGFWR 15
Db 86 DLGAYGFWSM 95

RESULT 11
Q9JX78 PRELIMINARY; PRT; 118 AA.
AC Q9JX78;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE NADH dehydrogenase I chain A [EC 1.6.5.3].
GN Name=nucA; OrderedLocusNames=NMA0019;
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919; DOI=10.1038/35006655;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltham T., Hamlin N., Holroyd S.,
RA Jagsels K., Leather S., Moule S., Mungall K.L., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis Z2491.";
RL Nature 404:502-506(2000).
CC -!- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
sulfur (Fe-S) centers, to quinones in the respiratory chain (By
similarity).
CC -!- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the complex I subunit 3 family.
DR EMBL; AL162752; CAB83339.1; -.
DR PIR; G81992; G81992.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.

```

Q25352 Q25352 PRELIMINARY; PRT; 188 AA.
AC Q25352; O2505;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE L2759-14.
GN Name=L2759.14;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=99178987; PubMed=10077609; DOI=10.1073/pnas.96.6.2902;
RA Myler P.J., Audleman L., DeVos T., Hixson G., Kiser P., Lemley C.,
RA Magness C., Rickel E., Sisk E., Sunkin S., Swartzell S., Westlake T.,
RA Bastien P., Fu G., Ivens A., Stuart K.;
RT "Leishmania major Friedlin chromosome 1 has an unusual distribution of
RT protein-coding genes";
RL Proc. Natl. Acad. Sci. U.S.A. 96:2902-2906(1999).
DR EMBL; AE001274; AAC24644.1; -;
DR PIR; B81459; T02822.
DR InterPro; IPR008506; DUF788.
DR Pfam; PF05620; DUF788; 1.
SQ SEQUENCE 188 AA; 20964 MW; EE4A9F2B776E8093B CRC64;
Query Match 49.4%; Score 43; DB 2; Length 188;
Best Local Similarity 77.8%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 HDLRAYGF 13
DB 46 HDLMAIGFW 54
RESULT 12
Q82MF7 Q82MF7 PRELIMINARY; PRT; 256 AA.
AC Q82MF7;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Putative proteoglycan 3,4-dioxygenase beta subunit.
GN Name=pcgH; OrderedLocNames=SAV1703;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Oonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005027; BAC69414.1; -;
DR HSSP; P00437; 3PCC.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008193; F:ferric iron binding; IEA.

DR GO; GO:0016702; F:oxidoreductase activity, acting on single d. . .; IEA.
DR GO; GO:0006725; P:aromatic compound metabolism; IEA.
DR InterPro; IPR000627; Dioxigenase.
DR Pfam; PF00775; Dioxigenase C; 1.
DR PROSITE; PS00083; INTRADIOL DIOXYGENAS; UNKNOWN 1.
KW Complete proteome; Dioxigenase.
SQ SEQUENCE 256 AA; 28537 MW; 73592DB9F6DAC8ED CRC64;
Query Match 49.4%; Score 43; DB 2; Length 256;
Best Local Similarity 72.7%; Pred. No. 46;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 SYQHDLRAYGF 12
DB 224 TYQHDLSAPGF 234
RESULT 13
Q6J925 Q6J925 PRELIMINARY; PRT; 430 AA.
AC Q6J925;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Recombination activating protein 1 (Fragment).
GN Name=RAG-1;
OS Arthroleptis variabilis.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Microhylidae;
OC Arthroleptidae; Arthroleptis.
OX NCBI_TaxID=83977;
RN [1]
RP SEQUENCE FROM N.A.
RA van der Meijden A., Vences M., Meyer A.;
RT "Novel Phylogenetic relationships of the enigmatic brevicipitine and
RT scaphiophrynine toads as revealed by sequences from the nuclear Rag-1
RT gene";
RL Proc. R. Soc. Lond., B, Biol. Sci. 0:0-0(2004).
DR EMBL; AY571642; AAT29041.1; -;
DR InterPro; IPR010988; M_synth_B12_like.
FT NON_TER 1
FT NON_TER 430
SQ SEQUENCE 430 AA; 49116 MW; 993256C18AF51AAE CRC64;
Query Match 49.4%; Score 43; DB 2; Length 430;
Best Local Similarity 50.0%; Pred. No. 78;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 RSYQHDLRAYGFWR 14
DB 234 RSHSENLRQYGTWR 247
RESULT 14
Q9JVV7 Q9JVV7 STANDARD; PRT; 459 AA.
AC Q9JVV7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Probable multidrug resistance protein norm (Na(+)/drug antiporter)
DE (Multidrug-efflux transporter).
GN Name=norm; OrderedLocNames=NMA1022;
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919; DOI=10.1038/35006655;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,

RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
 RA Jagels K., Leather S., Moule S., Mungall K.L., Quail M.A.,
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
 RA Whitehead S., Spratt B.G., Barrell B.G.,
 RT "Complete DNA sequence of a serogroup A strain of *Neisseria*
 RT meningitidis 22491.",
 RL Nature 404:502-506(2000).
 CC -!- FUNCTION: Functions as a Na(+)/drug antiporter (By similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (Potential).
 CC -!- SIMILARITY: Belongs to the multi antimicrobial extrusion (MATE)
 CC family. Norm (TC 2.A.66.1.1) subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, AL162754; CAB84291.1; -.
 DR PIR; D81950; D81950.
 DR HAMAP; MF_00400; -; 1.
 DR InterPro; IPR002528; Mate.
 DR Pfam; PF01554; Mate; 2.
 DR TIGRFAMs; TIGR00797; mate; 1.
 DR Antiport; Complete proteome; Inner membrane; Sodium transport;
 KW Transmembrane; Transport.
 FT TRANSMEM 20 40
 FT TRANSMEM 53 73
 FT TRANSMEM 100 120
 FT TRANSMEM 132 152
 FT TRANSMEM 168 188
 FT TRANSMEM 202 222
 FT TRANSMEM 252 272
 FT TRANSMEM 285 305
 FT TRANSMEM 325 345
 FT TRANSMEM 358 378
 FT TRANSMEM 395 415
 FT TRANSMEM 423 443
 FT TRANSMEM 459 494
 SQ SEQUENCE 459 AA; 49941 MW; B2EFD115C8EA494E CRC64;

Query Match 49.4%; Score 43; DB 1; Length 459;
 Best Local Similarity 50.0%; Pred. No. 83;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 SQYQHDRLRAYGFW 13
 :||:|:|:|:|
 Db 416 AYRFDWGIYGF 427

RESULT 15
 OS Bacillus stearothermophilus.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
 OX NCBI_TaxID=1422;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhu Z., Zhou J., Xu S.-Y.,
 RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY364324; AAQ72363.1; -.
 DR GO; GO:0004519; F:endonuclease activity; IEA.
 KW Endonuclease.
 SQ SEQUENCE 465 AA; 54697 MW; 3B72FD0F9AB3552C CRC64;

Query Match 49.4%; Score 43; DB 2; Length 465;
 Best Local Similarity 53.8%; Pred. No. 85;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 OY 1 RSYQHDRLRAYGFW 13
 :||:|:|:|:|
 Db 214 KSYTKDRRAVEYW 226
 :||:|:|:|:|
 RESULT 16
 OS Bacillus stearothermophilus.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
 OX NCBI_TaxID=1422;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Xu S.-Y., Dore A., Hume A., Pelletier J., Zhou J.,
 RT "Method for cloning and expression of BamBI restriction endonuclease
 RT and BamBI methylase in *E. coli* and purification of BamBI
 RT endonuclease.",
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY452068; AAS09914.1; -.
 DR GO; GO:0004519; F:endonuclease activity; IEA.
 KW Endonuclease.
 SQ SEQUENCE 530 AA; 62199 MW; 1443B3954B884B77 CRC64;

Query Match 49.4%; Score 43; DB 2; Length 530;
 Best Local Similarity 53.8%; Pred. No. 97;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 OY 1 RSYQHDRLRAYGFW 13
 :||:|:|:|:|
 Db 226 RLYNHDRRAFWW 238
 :||:|:|:|:|

RESULT 17
 OS Methanosarcina acetivorans.
 OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
 OX NCBI_TaxID=2214;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STAIN-C2A / ATCC 35395 / DSM 2834;
 RA MEDLINE=1932238; PubMed=11932238; DOI=10.1101/gr.223902;
 RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., MacDonald P.,
 RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atncor D., Brown A.,
 RA Allen N., Naylor J., Stange-Thomann N., DeArellano K., Johnson R.,
 RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
 RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
 RA Ferry J.G., Jarrell K.P., Jing H., Macario A.J.L., Paulsen I.T.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.
 RT "The genome of *Methanosarcina acetivorans* reveals extensive metabolic
 RT and physiological diversity."

Query Match 49.4%; Score 43; DB 2; Length 530;
 Best Local Similarity 53.8%; Pred. No. 97;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 OY 1 RSYQHDRLRAYGFW 13
 :||:|:|:|:|
 Db 226 RLYNHDRRAFWW 238
 :||:|:|:|:|


```

RL Genome Res. 12:532-542(2002).
DR EMBL; AE010884; AA053396.1; -.
KW Complete proteome.
SQ SEQUENCE 318 AA; 36371 MW; 377FBCF08D6576E0 CRC64;

Query Match 48.3%; Score 42; DB 2; Length 318;
Best Local Similarity 66.7%; Pred. No. 85;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 5 HDLRAYGF 13
Db 223 HELISYGF 231

RESULT 18
Q742P0 PRELIMINARY; PRT; 318 AA.
AC Q742P0;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MAP0795;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Anonin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017229; AA031112.1; -.
DR InterPro; IPR011251; Luciferase like.
DR Pfam; PF00296; Bac_luciferase; I.
KW Complete proteome.
SQ SEQUENCE 318 AA; 34907 MW; 867FE56C1817A5AF CRC64;

Query Match 48.3%; Score 42; DB 2; Length 318;
Best Local Similarity 60.0%; Pred. No. 85;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 YQHLRAYGF 12
Db 106 YEHEWYAYG 115

RESULT 19
Q67P30 PRELIMINARY; PRT; 321 AA.
AC Q67P30;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=STH1578;
OS Symbiobacterium thermophilum.
OC Bacteria; Actinobacteria; Symbiobacterium.
OX NCBI_TaxID=2734;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IAM14863;
RA Ueda K., Yamashita A., Ishikawa J., Shimada M., Watsuji T.,
RA Morimura K., Ikeda H., Hattori M., Beppu T.;
RT "Complete genome sequence of an uncultured bacterium Symbiobacterium
RT thermophilum."
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF006840; BAD40563.1; -.
KW Hypothetical protein.
SQ SEQUENCE 321 AA; 35228 MW; D974E5E125135E2B CRC64;

Query Match 48.3%; Score 42; DB 2; Length 321;
Best Local Similarity 63.6%; Pred. No. 86;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 QHDLRAYGF 14
Db 200 QHGRQAYQF 210

RESULT 20
Q8TRX2 PRELIMINARY; PRT; 418 AA.
AC Q8TRX2;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Predicted protein.
GN OrderedLocusNames=MA1043;
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238; DOI=10.1101/gr.223902;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.B., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.T.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity."
RL Genome Res. 12:532-542(2002).
DR EMBL; AE017770; AA04470.1; -.
DR SMART; SM00710; Pbh1; 8.
KW Complete proteome.
SQ SEQUENCE 418 AA; 47157 MW; B97C9A87C91A9563 CRC64;

Query Match 48.3%; Score 42; DB 2; Length 418;
Best Local Similarity 66.7%; Pred. No. 118+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 HDLRAYGF 13
Db 269 HDLCWGIW 277

RESULT 21
Q22405 PRELIMINARY; PRT; 444 AA.
AC Q22405;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein T11F9.1.
GN ORFNames=T11F9.1;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).

```

```

[2]
RN  SEQUENCE FROM N.A.
RP  STRAIN=Bristol N2;
RA  Lennard N.;
RL  Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RE  EMBL; 274042; CA98536.1; -.
DR  PIR; T24844; T24844.
DR  WormBase; WBGene0011709; T11F9.1.
DR  WormPep; T11F9.1; CE06413.
DR  GO; GO:0016021; C:integral to membrane; IEA.
DR  GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR  GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR  InterPro; IPR000276; GPCR_Rhodopsn.
DR  PRINTS; PR00237; GPCRHOOPS.
DR  PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.
KW  Hypothetical protein.
SQ  SEQUENCE 444 AA; 51686 MW; 3C07F095C5484892 CRC64;

Query Match      48.3%; Score 42; DB 2; Length 444;
Best Local Similarity 42.9%; Pred. No. 1.2e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 6; Gaps 1;

QY  1 RSYQHDLR-----YGFWR 15
Db  189 REYQLNLSALARYEYGYWRV 209

RESULT 22
QYVJHO PRELIMINARY; PRT; 652 AA.
AC  Q9VJHO;
DT  01-MAY-2000 (TrEMBLrel. 13, Created)
DT  01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT  01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE  CG13283-PA.
GN  ORFNames=CG13283;
OS  Drosophila melanogaster (Fruit fly).
OC  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC  Ephydroidea; Drosophilidae; Drosophila.
OX  NCBI_taxid=7227;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA  Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA  Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA  George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA  Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA  Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA  Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA  Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA  Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA  Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattier P.,
RA  Burris K.C., Buesam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA  Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA  de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA  Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA  Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA  Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA  Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA  Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA  Hootin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA  Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA  Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA  Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA  Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA  Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA  Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA  Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.M.,
RA  Palazzolo M., Pittman G.S., Pan S., Pollard J.R., Puri V., Reese M.G.,
RA  Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA  Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

```

```

RA  Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA  Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA  Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA  Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA  Yeh R.P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA  Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA  Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT  "The genome sequence of Drosophila melanogaster."
RL  Science 287:2185-2195(2000).
[2]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=22426065; PubMed=12537568;
RA  Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA  Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA  George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA  Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA  Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA  Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT  "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT  melanogaster euchromatic genome sequence."
RL  Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
[3]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=22426070; PubMed=12537573;
RA  Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA  Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA  Ashburner M., Celniker S.E.;
RT  "The transposable elements of the Drosophila melanogaster euchromatin:
RT  a genomics perspective."
RL  Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
[4]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=22426069; PubMed=12537572;
RA  Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA  Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA  Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA  Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA  Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA  Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA  Lewis S.E.;
RT  "Annotation of the Drosophila melanogaster euchromatic genome: a
RT  systematic review."
RL  Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
[5]
RP  SEQUENCE FROM N.A.
RG  FlyBase;
RL  Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
[6]
RP  SEQUENCE FROM N.A.
RG  FlyBase;
RL  Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AE003652; AAF53579.1; -.
DR  FlyBase; FBgn0032613; CG13283.
SQ  SEQUENCE 652 AA; 77836 MW; 8D4D549AB2D955A2 CRC64;

Query Match      48.3%; Score 42; DB 2; Length 652;
Best Local Similarity 53.8%; Pred. No. 1.8e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY  1 RSYQHDLRAYGFW 13
Db  557 RSFYSLTDYHFW 569

RESULT 23
SYA_BUCBP
ID  SYA_BUCBP STANDARD; PRT; 880 AA.
AC  P59420;
DT  10-OCT-2003 (Rel. 42, Created)
DT  10-OCT-2003 (Rel. 42, Last sequence update)
DT  25-OCT-2004 (Rel. 45, Last annotation update)
DE  Alanine--tRNA synthetase (EC 6.1.1.7) (Alanine--tRNA ligase) (AlARS).
GN  Name=alas; OrderedLocusNames=bbp364;

```



```

DR InterPro; IPR003938; EAG_ELK_ERG.
DR InterPro; IPR005821; Ion_trans.
DR InterPro; IPR001622; K+channel_pore.
DR InterPro; IPR005820; M+channel_nlg.
DR Pfam; PF00023; Ank; 5.
DR Pfam; PF00027; CNMP_binding; 1.
DR Pfam; PF00520; Ion_trans; 1.
DR PRINTS; PR01415; ANKYRIN.
DR PRINTS; PR01463; EAGCHANLFMLY.
DR SMART; SM00248; ANK; 4.
DR SMART; SM00100; CNMP; 1.
DR PROSITE; PS50068; ANK_REPEAT; 3.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50042; CNMP_BINDING; 3; 1.
KW ANK repeat; Ion transport; Ionic channel; Transmembrane; Transport.
SQ SEQUENCE 907 AA; 99461 MW; A0D9BAB721BD5E6E CRC64;

Query Match 48.3%; Score 42; DB 2; Length 907;
Best Local Similarity 56.2%; Pred. No. 2.5e+02;
Matches 9; Conservative 1; Mismatches 0; Indels 6; Gaps 1;

QY 6 DLRAYGF-----WRL 15
DB 169 DLRSYGFGLRLWRL 184

RESULT 26
Q7MUS9 PRELIMINARY; PRT; 951 AA.
AC Q7MUS9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=PG1405;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (Class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=W83.
RX MEDLINE=22829867; PubMed=12949112;
RX DOI=10.1128/JB.185.18.5591-5601.2003;
RA Nelson K.E., Fleischmann R.D., Deboy R.T., Paulsen I.T., Fouts D.E.,
RA Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,
RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
RA Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,
RA Dewhirst F.E., Fraser C.M.;
RT "Complete genome sequence of the oral pathogenic bacterium
RT Porphyromonas gingivalis strain W83.",
RL J. Bacteriol. 185:5591-5601(2003).
DR EMBL; AE017176; AAQ66463.1; --
DR TIGR; PG1405; --
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 951 AA; 108033 MW; D359CABA51C3C7B CRC64;

Query Match 48.3%; Score 42; DB 2; Length 951;
Best Local Similarity 56.2%; Pred. No. 2.6e+02;
Matches 9; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 2 SYQHDL--RAYGFWR 15
DB 599 SYMPDFTKRYGFWR 614

RESULT 27
SYLM_NEUCR
ID SYLM_NEUCR STANDARD; PRT; 994 AA.
AC P15181; Q7RVCS;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)

```

```

DE Leucyl-tRNA synthetase, mitochondrial precursor (EC 6.1.1.4)
DE (Leucine--tRNA ligase) (LeuRS).
GN Name=leu-5; ORFNames=99H12.170, NCU03814.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN SEQUENCE FROM N.A.
RX MEDLINE=90097874; PubMed=2574823;
RA Chow C.M., Metzger R.L., Rajbhandary U.L.;
RT "Nuclear gene for mitochondrial leucyl-tRNA synthetase of Neurospora
RT crassa: isolation, sequence, chromosomal mapping, and evidence that
RT the leu-5 locus specifies structural information.";
RL Mol. Cell. Biol. 9:4631-4644(1989).
RN [2]
RN SEQUENCE FROM N.A.
RX STRAIN=74-OR23-1A / FGSC 987;
RX MEDLINE=22542210; PubMed=12655011; DOI=10.1093/nar/gkg293;
RA Mannhaupt G., Montrone C., Haase D., Mewes H.-W., Aign V.,
RA Hoheisel J.D., Fartmann B., Nyakatura G., Kempken F., Maier J.,
RA Schulte U.;
RT "What's in the genome of a filamentous fungus? Analysis of the
RT Neurospora genome sequence.";
RL Nucleic Acids Res. 31:1944-1954(2003).
RN [3]
RN SEQUENCE FROM N.A.
RX STRAIN=74-OR23-1A / FGSC 987;
RX PubMed=12712197; DOI=10.1038/nature01554;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehan B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
Qui D., Tanakiev P., Bell-Pedersen D., Nelson M.A.,
RA Werner-Washburne M., Selitrenikoff C.P., Kinsey J.A., Braun E.L.,
RA Zelter A., Schulte U., Kothe G.O., Jedd G., Mewes H.-W., Staben C.,
RA Marcotte E., Greenberg D., Roy A., Foley K., Naylor J.,
RA Stange-Thomann N., Barrett R., Gnerre S., Kamal M., Kamysseis M.,
RA Mauceli E., Bielke C., Rudd S., Frishman D., Krystofova S.,
RA Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S., Cogoni C.,
RA Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA DeSouza C.P., Glass L., Orbach M.J., Berglund J.A., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT "The genome sequence of the filamentous fungus Neurospora crassa.";
RL Nature 422:859-868(2003)
CC -1- CATALYTIC ACTIVITY: ATP + L-leucine + tRNA (Leu) = AMP +
CC diphosphate + L-leucyl-tRNA(Leu).
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- SIMILARITY: Belongs to the class-I aminoacyl-tRNA synthetase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M30472; AAA33599.1; --
CC EMBL; AL451018; CAC18253.1; --
CC EMBL; AABX01000273; EAA31967.1; --
CC PIR; A33474; SYNCLM.
CC InterPro; IPR002302; Leu-tRNA-synt1a.
CC InterPro; IPR002300; tRNA-synt_1a.
CC InterPro; IPR001412; tRNA-synt_1.
CC InterPro; IPR009080; tRNA-synt_1a_bind.
CC InterPro; IPR009008; ValRS_IleRS_edit.
CC Pfam; PF00133; tRNA-synt_1; 1.
CC PRINTS; PR00985; TRNASYNTHLEU
CC TIGRFAMs; TIGR00396; leuS_bact; 1.
CC PROSITE; PS00176; AA_TRNA_LIGASE_I; 1.

```

KW Aminoacyl-tRNA synthetase; ATP-binding; Ligase; Mitochondrion;
 KW Protein biosynthesis; Transit peptide.
 FT TRANSIT 1 52 Mitochondrion.
 FT CHAIN 53 994 Leucyl-tRNA synthetase.
 FT SITE 93 103 "HIGH" region.
 FT SITE 697 701 "KMSKS" region.
 FT BINDING 700 700 ATP (by similarity).
 SQ SEQUENCE 994 AA; 109853 MW; 2ADB26302A8465F9 CRC64;

Query Match 48.3%; Score 42; DB 1; Length 994;
 Best Local Similarity 60.0%; Pred. No. 2.8e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 HDLRAYGFW 14
 ||||| : |||
 DB 403 HDLRDHAFWK 412

RESULT 28

Q9NEG4 PRELIMINARY; PRT; 306 AA.
 AC Q9NEG4
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein L2954.12.
 GN Name=L2954.12;
 OS Leishmania major.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Friedlin;
 RX MEDLINE=98146435; PubMed=9477341;
 RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
 RA Smith D.F.;
 RT "A physical map of the Leishmania major Friedlin genome."
 RL Genome Res. 8:135-145(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Friedlin;
 RA Borzym K., Klages S., Reinhardt R., Beck A., Ivens A.C., Quail M.,
 RA Rajandream M.A., Barrell B.G.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL138618; CAB71284.2; --
 KW Hypothetical protein.
 SQ SEQUENCE 306 AA; 34129 MW; D03DA12477200AE7 CRC64;

Query Match 47.7%; Score 41.5; DB 2; Length 306;
 Best Local Similarity 66.7%; Pred. No. 99;
 Matches 8; Conservative 1; Mismatches 0; Indels 3; Gaps 1;

QY 4 QHDLRAYGFWRL 15
 ||| : |||||
 DB 213 QHN---YGFWR 221

RESULT 29

Q721X3 PRELIMINARY; PRT; 60 AA.
 AC Q721X3
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Alpha-2-macroglobulin family (Fragment).
 OS Botryllus schlosseri (Star ascidian).
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
 OC Stolidobranchia; Botryllidae; Botryllus.
 OX NCBI_TaxID=30301;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bartl S., Baish M., Weisman I.L., Diaz M.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY102167; AAM52096.1; --
 DR GO; GO:0004866; F:endorpeptidase inhibitor activity; IEA.
 DR InterPro; IPR008930; Terp_cyc_toroid.
 FT NON_TER 1 60
 FT SITE 60 60
 SQ SEQUENCE 60 AA; 6911 MW; 9637A853A698D685 CRC64;

Query Match 47.1%; Score 41; DB 2; Length 60;
 Best Local Similarity 46.2%; Pred. No. 22;
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 RSYQHDLRAYGFW 13
 ::||: |||
 DB 39 QNYRHDNGAYSIV 51

RESULT 30

Q9H356 PRELIMINARY; PRT; 61 AA.
 AC Q9H356
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE PRO1804.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Zhang C., Yu Y., Zhang S., Zhou G., Wei H., Bi J., Dong C., Zai Y.,
 RA Xu W., Gao F., Liu M., He P.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF132201; AAG35547.1; --
 SQ SEQUENCE 61 AA; 6678 MW; F238B0E082E080C9 CRC64;

Query Match 47.1%; Score 41; DB 2; Length 61;
 Best Local Similarity 58.3%; Pred. No. 23;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 SYQHDLRAYGFW 13
 ||| : |||||
 DB 6 SYQNSLLGLGFW 17

RESULT 31

Q8D683 PRELIMINARY; PRT; 61 AA.
 AC Q8D683
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Transposase.
 GN OrderedLocusNames=VV20654;
 OS Vibrio vulnificus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=672;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CMCP6;
 RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
 RA Choy H.E.;
 RT "Complete genome sequence of Vibrio vulnificus CMCP6."
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE016810; AA007596.1; --
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006310; P:DNA recombination; IEA.
 DR InterPro; IPR001584; Rve.
 DR Pfam; PF00665; rve; 1.
 KW Complete proteome.
 SQ SEQUENCE 61 AA; 7115 MW; A424640F25849FA7 CRC64;

4

ID Q8PW04 PRELIMINARY; PRT; 229 AA.
AC Q8PW04;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Transporter.
GN OrderedLocustNames=WM1800;
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goel / GoI / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wierse A., Baeumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boencke M., Steckel S.,
RA Bhattacharya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosarcina mazei: evidence for lateral gene
RT transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR EMBL: AE013416; AAM31496.1; --
DR InterPro: IPR003744; DUF165.
DR Pfam: PF02592; DUF165; 1.
DR TIGRFAMs: TIGR00697; DUF165; 1.
KW Complete proteome.
SQ SEQUENCE 229 AA; 25476 MW; D5495B8192985291 CRC64;

Query Match 47.1%; Score 41; DB 2; Length 229;
Best Local Similarity 46.2%; Pred. No. 89;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 SYQHDRLRAYGFW 14
DB 110 SYQHDIAWAFNFK 122

RESULT 36
Q8G977 PRELIMINARY; PRT; 260 AA.
AC Q8G977;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative transposase (Fragment).
GN Name=tnpA;
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PP3;
RX MEDLINE=22313464; PubMed=12426347;
RX DOI=10.1128/JB.184.23.6581-6591.2002;
RA Weightman A.J., Topping A.W., Hill K.E., Lee L.;
RT "Investigation of two evolutionarily unrelated halocarboxylic acid
RT dehalogenase gene families";
RL J. Bacteriol. 184:6581-6591(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PP3;
RA Hill H.K., Weightman A.J.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ534881; CAD59132.1; --
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0004803; F:transposase activity; IEA.
DR GO: GO:0006310; P:DNA recombination; IEA.
DR InterPro: IPR002560; Transposase_12.
DR Pfam: PF01610; Transposase_12; 1.
NON_TER 1

FT NON_TER 260 260
SQ SEQUENCE 260 AA; 30910 MW; 32BC1PA00D6ABF3B CRC64;

Query Match 47.1%; Score 41; DB 2; Length 260;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 SYQHDRLRAYGFW 13
DB 228 AYEHKRFYGIW 239

RESULT 37
Q8GBV1 PRELIMINARY; PRT; 260 AA.
AC Q8GBV1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative transposase (Fragment).
GN Name=tnpA;
OS Pseudomonas sp. F12.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=217730;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F12;
RA Hill K.E., Weightman A.J.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ534880; CAD59131.1; --
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0004803; F:transposase activity; IEA.
DR GO: GO:0006310; P:DNA recombination; IEA.
DR InterPro: IPR002560; Transposase_12.
DR Pfam: PF01610; Transposase_12; 1.
NON_TER 1
FT NON_TER 260 260
SQ SEQUENCE 260 AA; 30910 MW; 32BC1PA00D6ABF3B CRC64;

Query Match 47.1%; Score 41; DB 2; Length 260;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 SYQHDRLRAYGFW 13
DB 228 AYEHKRFYGIW 239

RESULT 38
RL5_NEUCR STANDARD; PRT; 301 AA.
AC O59553; Q7RVB8;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE 60S ribosomal protein L5 (CPR4).
GN Name=rlp-5; Synonyms=crp-4; ORFNames=B7H23.350, NCU04331.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A / FGSC 987;
RA de la Serna I.L., Cujec T.P., Shi Y., Tyler B.M.;
RT "Non-coordinate regulation of 5S rRNA genes and the gene encoding the
RT 5S rRNA-binding ribosomal protein homolog in Neurospora crassa.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A / FGSC 987;
RX MEDLINE=22542210; PubMed=12655011; DOI=10.1093/nar/gk9293;

Manhaupt G., Montrone C., Haase D., Mewes H.-W., Aign V.,
Honeisei J.D., Fartmann B., Nyakatura G., Kempken F., Maier J.,
Schulte U.;
"What's in the genome of a filamentous fungus? Analysis of the
Neurospora genome sequence.";
Nucleic Acids Res. 31:1944-1954 (2003).
[3]
SEQUENCE FROM N.A.
RP STRAIN=74-OR23-1A / FGSC 987;
RX PubMed=12712197; DOI=10.1038/nature01554;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Reiman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Tanakiev P., Bell-Pedersen D., Nelson M.A.,
RA Werner-Washburne P., Selitrennikoff C.P., Kinsey J.A., Braun E.L.,
RA Zelter A., Schulte U., Kothe G.O., Jedd G., Mewes H.-W., Staben C.,
RA Marcotte E., Greenberg D., Roy A., Foley K., Naylor J.,
RA Stange-Thomann N., Barrett R., Gnerre S., Kamal M., Kamysov M.,
RA Mauceli E., Bielke C., Rudd S., Frisman D., Krystofova S.,
RA Magnusen C., Metzgen R.L., Perkins D.D., Kroken S., Cogoni C.,
RA Racino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA Desouza C.P., Glass L., Orbach M.J., Berglund J.A., Voelker R.,
RA Yarden O., Plamann M., Sella S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannheim G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT "The genome sequence of the filamentous fungus Neurospora crassa.";
RL Nature 422:859-868 (2003).
CC -1- FUNCTION: This protein binds 5S RNA (By similarity).
CC -1- SIMILARITY: Belongs to the ribosomal protein L18p family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF054907; AAC09000.1; -;
DR EMBL; BX294026; CAD71058.1; -;
DR EMBL; ABX01000314; EAA31342.1; -;
DR InterPro; IPR005484; Ribosomal_L18p.
DR InterPro; IPR005485; Ribosomal_L5euk.
DR Pfam; PF00861; Ribosomal_L18p; 1.
DR PRINTS; PR00058; RIBOSOMALLS.
DR ProDom; PD001394; Ribosomal_L18p; 1.
KW Ribosomal protein; rRNA-binding.
SQ SEQUENCE 301 AA; 34411 MW; 289B0F0D1672A5F9 CRC64;

Query Match 47.1%; Score 41; DB 1; Length 301;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 SYQHDLRAYG 11
DB :||:|:
78 AYSHELKAYG 87

RESULT 39
QG6MAL4 PRELIMINARY; PRT; 323 AA.
AC QG6MAL4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=pci1661;
OS Parachlamydia sp. (strain UWE25) (subsp. Acanthamoeba sp.).
OC Bacteria; Chlamydiae; Chlamydiales; Parachlamydiaceae; Parachlamydia.
OC NCBI_TaxID=264201;
RN [1]
RP SEQUENCE FROM N.A.
RA Horn M., Collingro A., Schmitz-Esser S., Beier C.L., Purkhold U.,

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 18:31:37 ; Search time 164 Seconds
(without alignments)
35.374 Million cell updates/sec

Title: US-09-623-038-6

Perfect score: 15

Sequence: 1 RSYQHDLRAYGFWR 15

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 6

Total number of hits satisfying chosen parameters: 35

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : A.Geneseq 16Dec04:*

- 1: Genesep1980s:*
- 2: Genesep1990s:*
- 3: Genesep2000s:*
- 4: Genesep2001s:*
- 5: Genesep2002s:*
- 6: Genesep2003as:*
- 7: Genesep2003bs:*
- 8: Genesep2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 15 | 100.0 | 15 | 2 AAY30352 | Aay30352 Epitope d |
| 2 | 15 | 100.0 | 15 | 5 AAE19240 | Aae19240 Streptoco |
| 3 | 15 | 100.0 | 66 | 5 AAE19245 | Aae19245 Bi-peptid |
| 4 | 14 | 93.3 | 15 | 5 AAE19244 | Aae19244 Streptoco |
| 5 | 14 | 93.3 | 47 | 5 AAE19247 | Aae19247 Tri-pepti |
| 6 | 6 | 40.0 | 12 | 2 AAR14929 | Aar14929 OT-2 anti |
| 7 | 6 | 40.0 | 143 | 2 AAW71540 | Aaw71540 Helicobac |
| 8 | 6 | 40.0 | 144 | 7 ADF05142 | Adf05142 Bacterial |
| 9 | 6 | 40.0 | 187 | 2 AAW20436 | Aaw20436 H. pylori |
| 10 | 6 | 40.0 | 428 | 8 ADS27611 | Ads27611 Bacterial |
| 11 | 6 | 40.0 | 432 | 2 AAW20733 | Aaw20733 H. pylori |
| 12 | 6 | 40.0 | 447 | 4 AAU36020 | Aau36020 Helicobac |
| 13 | 6 | 40.0 | 466 | 6 AEU32026 | Aeu32026 Protein e |
| 14 | 6 | 40.0 | 476 | 7 ADC08077 | Adc08077 Rice prot |
| 15 | 6 | 40.0 | 484 | 7 ABO63378 | Abo63378 Klebsiell |
| 16 | 6 | 40.0 | 515 | 4 AAU02903 | Aau02903 Angiotens |
| 17 | 6 | 40.0 | 615 | 4 AAM40056 | Aam40056 Human pol |
| 18 | 6 | 40.0 | 615 | 7 ADA50545 | Ada50545 Human fac |
| 19 | 6 | 40.0 | 615 | 8 ADN04180 | Adn04180 Antipsori |
| 20 | 6 | 40.0 | 615 | 8 ADM98382 | Adm98382 Human fac |
| 21 | 6 | 40.0 | 615 | 8 ABM80709 | Abm80709 Tumour-as |
| 22 | 6 | 40.0 | 621 | 5 AAU11271 | Aau11271 Human coa |
| 23 | 6 | 40.0 | 621 | 5 AAU11270 | Aau11270 Human coa |
| 24 | 6 | 40.0 | 631 | 8 ADI32366 | Adi32366 Rice phos |
| 25 | 6 | 40.0 | 653 | 6 ABB82636 | Abb82636 PERV enve |

26 6 40.0 653 6 ABB82635
27 6 40.0 657 2 AAW32095
28 6 40.0 657 8 ADS73445
29 6 40.0 659 4 AAB35112
30 6 40.0 660 2 AAW85453
31 6 40.0 660 4 AAB35113
32 6 40.0 660 4 AAB70632
33 6 40.0 1069 4 AAB73284
34 6 40.0 1084 8 ADN73089
35 6 40.0 1420 8 ADR86081

ALIGNMENTS

RESULT 1

AAV30352
ID AAY30352 standard; peptide; 15 AA.

XX
AC AAY30352;

XX
DT 09-NOV-1999 (first entry)

XX
DE Epitope derived from pneumococcal surface adhesion A protein.

XX
KW Pneumococcal surface adhesion A protein; PsaA; monoclonal antibody;

XX
KW vaccine; Streptococcus pneumoniae infection.

XX
OS Streptococcus pneumoniae.

XX
FN WO9945121-A1.

XX
PD 10-SEP-1999.

XX
PF 26-FEB-1999; 99WO-US004326.

XX
PR 02-MAR-1998; 98US-0076565P.

XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX
PI Carlone GM, Ades EW, Sampson JS, Tharpe JA, Zeiler JL;

PI
Westerink MAJ;

XX
DR WPI; 1999-540849/45.

XX
PT New peptides corresponding to Streptococcus pneumoniae PsaA, used for treating or preventing Streptococcus pneumoniae infection in a subject.

XX
PS Claim 6; Page 43; 58pp; English.

XX
CC AAY30351-54 represent immunogenic peptides which are derived from a pneumococcal surface adhesion A protein (PsaA). The specification describes monoclonal antibodies which bind epitopes of the PsaA protein (e.g present sequence). The peptides can be used in vaccines to prevent Streptococcus pneumoniae infections. The antibodies of the invention can also be used to detect S. pneumoniae in a sample or individual

XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 15; DB 2; Length 15;

Best Local Similarity 100.0%; Pred. No. 2.5e-09;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RSYQHDLRAYGFWR 15

DB 1 RSYQHDLRAYGFWR 15

RESULT 2

AAE19240

ID AAE19240 standard; peptide; 15 AA.

XX
AC AAE19240;

```
XX 21-MAY-2002 (first entry)
XX Streptococcus pneumoniae PsaA immunogenic peptide #2.
XX Multiple antigenic peptide; MAP; immunogenic; immunity; infection;
XX pneumococcal surface adhesin protein A; PsaA; antibacterial.
XX Streptococcus pneumoniae.
XX WO200204497-A2.
XX 17-JAN-2002.
XX 10-JUL-2001; 2001WO-US021626.
XX 10-JUL-2000; 2000US-00613092.
XX 10-JUL-2000; 2000US-00613092.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX Ades EW, Johnson SE, Jue DL, Sampson JS, Carlone GM;
XX WPI; 2002-195762/25.
XX New multiple antigenic peptide for immunizing against streptococcal
XX infections, binds to monoclonal antibody obtained in response to
XX immunizing an animal with pneumococcal surface adhesion protein A or its
XX fragment.
XX Claim 2; Page 56; 86pp; English.
XX The invention relates to multiple antigenic peptides (MAP) immunogenic
XX against Streptococcus pneumoniae. MAP binds to monoclonal antibody
XX obtained in response to immunising an animal with pneumococcal surface
XX adhesion protein A (PsaA) or its fragment. MAP is useful for conferring
XX protective immunity against S. pneumoniae infection in a subject. The
XX present sequence is a bi-peptide heterogenous MAP used to provide
XX protection against Streptococcus pneumoniae challenge
XX Sequence 15 AA;
XX
Query Match 100.0%; Score 15; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.5e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 RSYQHDLRAYGFWR 15
Db 1 RSYQHDLRAYGFWR 15
RESULT 3
AAE19245
ID AAE19245 standard; protein; 66 AA.
XX AAE19245;
XX 21-MAY-2002 (first entry)
XX Bi-peptide heterogenous multiple antigenic peptide (MAP) #1.
XX Multiple antigenic peptide; MAP; immunogenic; immunity; infection;
XX pneumococcal surface adhesin protein A; PsaA; antibacterial.
XX Unidentified.
XX Location/Qualifiers
Key 1..15
Peptide /note= "PsaA immunogenic peptide #1"
FT 16..30
Peptide /note= "PsaA immunogenic peptide #2"
FT 31..45
Peptide /note= "PsaA immunogenic peptide #1"
FT 46..60
Peptide /note= "PsaA immunogenic peptide #2"
```

```
FT Modified-site 62
FT /label= Nle
FT Modified-site 64
FT /label= Nle
FT Modified-site 66
FT /label= Nle
XX WO200204497-A2.
XX 17-JAN-2002.
XX 10-JUL-2001; 2001WO-US021626.
XX 10-JUL-2000; 2000US-00613092.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX Ades EW, Johnson SE, Jue DL, Sampson JS, Carlone GM;
XX WPI; 2002-195762/25.
XX New multiple antigenic peptide for immunizing against streptococcal
XX infections, binds to monoclonal antibody obtained in response to
XX immunizing an animal with pneumococcal surface adhesion protein A or its
XX fragment.
XX Example 1b; Fig 1A; 86pp; English.
XX The invention relates to multiple antigenic peptides (MAP) immunogenic
XX against Streptococcus pneumoniae. MAP binds to monoclonal antibody
XX obtained in response to immunising an animal with pneumococcal surface
XX adhesion protein A (PsaA) or its fragment. MAP is useful for conferring
XX protective immunity against S. pneumoniae infection in a subject. The
XX present sequence is a bi-peptide heterogenous MAP used to provide
XX protection against Streptococcus pneumoniae challenge
XX Sequence 66 AA;
XX
Query Match 100.0%; Score 15; DB 5; Length 66;
Best Local Similarity 100.0%; Pred. No. 8.6e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 RSYQHDLRAYGFWR 15
Db 16 RSYQHDLRAYGFWR 30
RESULT 4
AAE19244
ID AAE19244 standard; peptide; 15 AA.
XX AAE19244;
XX 21-MAY-2002 (first entry)
XX Streptococcus pneumoniae PsaA immunogenic peptide #6.
XX Multiple antigenic peptide; MAP; immunogenic; immunity; infection;
XX pneumococcal surface adhesin protein A; PsaA; antibacterial.
XX Streptococcus pneumoniae.
XX WO200204497-A2.
XX 17-JAN-2002.
XX 10-JUL-2001; 2001WO-US021626.
XX 10-JUL-2000; 2000US-00613092.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX Ades EW, Johnson SE, Jue DL, Sampson JS, Carlone GM;
```

```

XX WPI; 2002-195762/25.
XX
XX New multiple antigenic peptide for immunizing against streptococcal
XX infections, binds to monoclonal antibody obtained in response to
XX immunizing an animal with pneumococcal surface adhesion protein A or its
XX fragment.
XX
XX Claim 2; Page 86; 86pp; English.
XX
XX The invention relates to multiple antigenic peptides (MAP) immunogenic
XX against Streptococcus pneumoniae. MAP binds to monoclonal antibody
XX obtained in response to immunising an animal with pneumococcal surface
XX adhesion protein A (PsaA) or its fragment. MAP is useful for conferring
XX protective immunity against S. pneumoniae infection in a subject. The
XX present sequence is Streptococcus pneumoniae PsaA immunogenic peptide
XX
XX Sequence 15 AA;
XX
XX Query Match 93.3%; Score 14; DB 5; Length 15;
XX Best Local Similarity 100.0%; Pred. No. 3e-08;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 SYQHDRLRAYGFWRLL 15
XX |||||
XX DB 1 SYQHDRLRAYGFWRLL 14
XX
XX RESULT 5
XX AAEL19247
XX ID AAEL19247 standard; protein; 47 AA.
XX
XX AC AAEL19247;
XX
XX DT 21-MAY-2002 (first entry)
XX
XX DE Tri-peptide heterogenous multiple antigenic peptide (MAP).
XX
XX KW Multiple antigenic peptide; MAP; immunogenic; immunity; infection;
XX pneumococcal surface adhesion protein A; PsaA; antibacterial.
XX
XX OS Unidentified.
XX
XX FH Key Location/Qualifiers
XX FT Peptide 1..15
XX FT /note= "PsaA immunogenic peptide #1"
XX FT Peptide 16..29
XX FT /note= "PsaA immunogenic peptide #6"
XX FT Peptide 30..43
XX FT /note= "PsaA immunogenic peptide #5"
XX FT Modified-site 45
XX FT /label= Nle
XX FT Modified-site 47
XX FT /label= Nle
XX
XX PN WO200204497-A2.
XX
XX PD 17-JAN-2002.
XX
XX PF 10-JUL-2001; 2001WO-US021626.
XX
XX PR 10-JUL-2000; 2000US-00613092.
XX
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX PI Ades EW, Johnson SE, Jue DL, Sampson JS, Carlone GW;
XX
XX WPI; 2002-195762/25.
XX
XX New multiple antigenic peptide for immunizing against streptococcal
XX infections, binds to monoclonal antibody obtained in response to
XX immunizing an animal with pneumococcal surface adhesion protein A or its
XX fragment.
XX
XX Example 1b; Fig 1C; 86pp; English.
XX
XX The invention relates to multiple antigenic peptides (MAP) immunogenic
XX against Streptococcus pneumoniae. MAP binds to monoclonal antibody
XX obtained in response to immunising an animal with pneumococcal surface
XX adhesion protein A (PsaA) or its fragment. MAP is useful for conferring
XX protective immunity against S. pneumoniae infection in a subject. The
XX present sequence is a tri-peptide heterogenous MAP used to provide
XX protection against Streptococcus pneumoniae challenge
XX
XX Sequence 47 AA;
XX
XX Query Match 93.3%; Score 14; DB 5; Length 47;
XX Best Local Similarity 100.0%; Pred. No. 7.6e-08;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 SYQHDRLRAYGFWRLL 15
XX |||||
XX DB 16 SYQHDRLRAYGFWRLL 29
XX
XX RESULT 6
XX AAR14929
XX ID AAR14929 standard; protein; 12 AA.
XX
XX AC AAR14929;
XX
XX DT 13-FEB-1992 (first entry)
XX
XX DE OT-2 antibody binding peptide (2).
XX
XX KW Monoclonal antibody; antigen; immunogen; Factor XII; epitope.
XX
XX OS Synthetic.
XX
XX FN WO9117258-A.
XX
XX PD 14-NOV-1991.
XX
XX PF 10-MAY-1990; 90US-00521820.
XX
XX PR 10-MAY-1990; 90US-00521820.
XX
XX PA (CETU ) CETUS CORP.
XX
XX PI Nuijens JH, Huijbregts CCM, Hack CE;
XX
XX WPI; 1991-353779/48.
XX
XX Treatment of sepsis using inhibitor of factor XII activation - comprises
XX use of new OT-2 antibody.
XX
XX Claim 15,17; Page 24; 32pp; English.
XX
XX Based on the known amino acid sequence of Factor XII, peptides corresp.
XX to neutralising epitopes of the mol. are synthesised and used as
XX immunogens to produce antibody. The pref. peptides are represented in
XX AAR14928-30. Amino acid Asp in this sequence = residue 442
XX
XX Sequence 12 AA;
XX
XX Query Match 40.0%; Score 6; DB 2; Length 12;
XX Best Local Similarity 100.0%; Pred. No. 9.1;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 SYQHDLL 7
XX |||||
XX DB 5 SYQHDLL 10
XX
XX RESULT 7
XX AAW71540

```

ADP05142 standard; protein; 143 AA.

AAW71540;

09-NOV-1998 (first entry)

Helicobacter polypeptide GHPO 956.

GHPO 956; infection; therapy; diagnosis; vaccine; gastritis; ulcer.

Helicobacter pylori.

Key Location/Qualifiers
Peptide 1..31
/label= Sig_peptide
Protein 32..143
/label= Mat_protein

WO9821225-A1.

22-MAY-1998.

14-NOV-1997; 97WO-05021353.

14-NOV-1996; 96US-00749051.

01-APR-1997; 97US-00831309.

01-APR-1997; 97US-00833457.

01-APR-1997; 97US-00834705.

24-JUN-1997; 97US-00881227.

29-JUL-1997; 97US-00902615.

(INNR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.

(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

(HUMA-) HUMAN GENOME SCI INC.

Haas R, Kleanthous H, Tomb J, Miller C, Al-Garawi A, Odenbreit S, Meyer T;

WPI; 1998-297855/26.

N-PSDB; AAV52075.

Helicobacter polynucleotide and polypeptide sequences - useful to treat or prevent gastrointestinal infection.

Claim 1; Page 272; 362pp; English.

This claimed Helicobacter pylori polypeptide, designated GHPO 956, can be used in vaccination methods for preventing or treating Helicobacter infection. 85 Helicobacter polypeptides (see AAW71474-W71558) are claimed, as well as isolated polynucleotides (see AAV52009-93) that encode them. The invention also provides: methods for producing these Helicobacter polypeptides in recombinant host systems, and related expression cassettes, vectors and transformed or transfected host cells; live vaccine vectors that contain the polynucleotides of the invention and which can be used to prevent or treat Helicobacter infection; therapeutic and/or prophylactic methods involving administration of polynucleotide molecules, polypeptides or monospecific antibodies; methods for detecting the presence of Helicobacter in samples using e.g. the polypeptides or monospecific antibodies; and methods for purifying the polypeptides by antibody-based affinity chromatography

Sequence 143 AA;

Query Match 40.0%; Score 6; DB 2; Length 143;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LRAYGF 12

Db 74 LRAYGF 79

RESULT 8

ADP05142

ID ADF05142 standard; protein; 144 AA.

XX

AC ADF05142;

XX

DT 12-FEB-2004 (first entry)

XX

DE Bacterial polypeptide #1255.

XX

KW Proteus mirabilis infection; bacterial infection; antibacterial;

XX

OS Proteus mirabilis.

XX

PN US6605709-B1.

XX

PD 12-AUG-2003.

XX

PF 05-APR-2000; 2000US-00543681.

XX

PR 09-APR-1999; 99US-0128706P.

XX

PA (GENO-) GENOME THERAPEUTICS CORP.

XX

PI Breton GL;

XX

DR WPI; 2003-895291/82.

XX

DR N-PSDB; ADF00970.

XX

PT New Proteus mirabilis polypeptides and polynucleotides, useful as

XX

PS Disclosure; SEQ ID NO 5427; 870pp; English.

XX

The invention relates to new Proteus mirabilis polypeptides and polynucleotides. The invention also relates to antibodies against the polypeptides, methods for producing the polypeptides, a method of generating vaccines for immunising an individual against P. mirabilis, a method for evaluating a compound for the ability to bind a P. mirabilis polypeptide and a method for screening test compounds for anti-bacterial activity. The polypeptides and polynucleotides are useful as molecular targets for diagnosing, preventing and treating pathological conditions resulting from bacterial infection, as reagents for diagnosis of bacterial diseases, as components of antibacterial vaccines, as targets for antibacterial drugs or as bio-control agents for plants. This sequence represents a Proteus mirabilis polypeptide of the invention.

Sequence 144 AA;

Query Match 40.0%; Score 6; DB 7; Length 144;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 YQHDLR 8

Db 56 YQHDLR 61

RESULT 9

AAW20436

ID AAW20436 standard; protein; 187 AA.

XX

AC AAW20436;

XX

DT 14-JUL-1997 (first entry)

XX

DE H. pylori protein.

XX

KW Cytoplasmic; vaccine; prevention; treatment; infection; identification; binding compound; bacterium; life cycle; activator; bacteria; inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis; envelope;

KW outer membrane; cell envelope; transporter.
 XX Helicobacter pylori.
 OS WO9640893-A1.
 PN XX
 XX 19-DEC-1996.
 PD XX
 XX 06-JUN-1996; 96WO-US009122.
 PF XX
 XX 07-JUN-1995; 95US-00487032.
 PR XX
 XX 01-APR-1996; 96US-00630405.
 XX XX
 XX (ASTR) ASTRA AB.
 PA XX
 XX Smith D, Berglindh OT, Mellgaard BL;
 PI WPI; 1997-0522306/05.
 DR N-PSDB; AAT67609.
 XX Helicobacter pylori nucleic acid sequences and related polypeptide(s) -
 PT useful for vaccines to treat or prevent H. pylori infection, and to
 PT detect Helicobacter.
 PT XX
 XX Disclosure; Page 613; 1481pp; English.
 PS CC
 XX The present sequence is a Helicobacter pylori protein of unknown
 CC function. The protein may be used in a vaccine to prevent or treat H.
 CC pylori infection or to identify H. pylori polypeptide binding compounds,
 CC useful as potential H. pylori life cycle activators or inhibitors. The
 CC genomic sequence of H. pylori (ATCC 55679) was determined from
 CC overlapping contigs generated by mechanically shearing the bacterial DNA.
 CC The sequences were analysed for ORF of at least 180 nucleotides, and the
 CC predicted coding regions defined by computer evaluation. To identify
 CC likely H. pylori antigens for vaccine development, the amino acid
 CC sequences predicted from various ORF were analysed for significant
 CC homology to other known or exported membrane proteins. Having identified
 CC and determined the sequences of interest, particular regions can be
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide
 CC production, e.g. in E. coli hosts
 XX Sequence 187 AA;
 SQ
 Query Match 40.0%; Score 6; DB 2; Length 187;
 Best Local Similarity 100.0%; Pred. No. 87;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 YQHDLR 8
 Db 103 YQHDLR 108
 |||||
 RESULT 10
 ID ADS27611 standard; protein; 428 AA.
 XX
 XX ADS27611;
 AC
 XX 02-DEC-2004 (first entry)
 DT
 XX Bacterial polypeptide #16644.
 DE
 XX Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 XX bacterial polypeptide.
 KW Bacteria.
 XX
 XX US2003233675-A1.
 PN

XX 18-DEC-2003.
 PD
 XX 20-FEB-2003; 2003US-00369493.
 PF
 XX 21-FEB-2002; 2002US-0360039P.
 PR
 XX (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 PI WPI; 2004-061375/06.
 DR
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 PT
 XX Claim 1; SEQ ID NO 16644; 122pp; English.
 PS
 XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 XX Sequence 428 AA;
 SQ
 Query Match 40.0%; Score 6; DB 8; Length 428;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 LRAYGF 12
 Db 44 LRAYGF 49
 |||||
 RESULT 11
 ID AAW20733 standard; protein; 432 AA.
 XX
 XX AAW20733;
 AC
 XX 16-JUL-1997 (first entry)
 DT
 XX H. pylori cell envelope protein, 06cpl1722orf15.
 DE
 XX Cytoplasmic; vaccine; prevention; treatment; infection; identification;
 KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.
 XX

PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlson KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA35896.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 25; SEQ ID NO 59950; 1766pp; English.
 PS The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway;
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 466 AA;
 SQ Query Match 40.0%; Score 6; DB 6; Length 466;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 10 YGFWRLL 15
 Db 432 YGFWRLL 437
 RESULT 14
 ADC08077
 ID ADC08077 standard; protein; 476 AA.
 XX ADC08077;
 AC ADC08077;
 XX Rice protein sequence Seq ID344 related to grain filling.
 DT 18-DEC-2003 (first entry)
 XX plant biotechnology; carbohydrate synthesis; carbohydrate metabolism;
 XX carbohydrate degradation; carbohydrate; plant grain; grain filling; corn;
 KW tomato; banana; canola; cotton; sorghum; tobacco; sugarbeet;
 KW wheat; rice; protein; oil; starch; fibre; moisture content; cereal grain;

KW gene; ds; plant.
 XX Oryza sativa.
 OS WO2003000905-A2.
 PN 03-JAN-2003.
 XX 21-JUN-2002; 2002WO-IB002450.
 PF 22-JUN-2001; 2001US-0300112P.
 PR 26-SEP-2001; 2001US-0325277P.
 PR 20-DEC-2001; 2001US-0342327P.
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 PA Zhu T, Cheng W, Briggs S, Cooper B, Goff SA, Moughamer T;
 PI Glazebrook J, Katagiri F, Kreps J, Provart N, Rieke D;
 XX WPI; 2003-229341/22.
 DR N-PSDB; ADC08076.
 XX New plant genes encoding polypeptides having an activity involved in or
 PT associated with the synthesis, metabolism or degradation of carbohydrates
 PT in the plant grain useful in generating plants having improved
 PT nutritional properties.
 XX Claim 34; SEQ ID NO 344; 130pp; English.
 PS This invention, in the area of plant biotechnology, relates to novel
 CC polynucleotides comprising a nucleotide sequence encoding a protein which
 CC is involved in or associated with the synthesis, metabolism or
 CC degradation of carbohydrates in the plant grain and the expression of
 CC which is up-regulated during grain filling. The plant is selected from
 CC corn, tomato, banana, canola, cotton, peanut, sorghum, tobacco,
 CC sugarbeet, wheat, and rice. The invention may be useful for the
 CC improvement of protein, oil, starch, fibre and moisture content of the
 CC cereal grains. In addition, carbohydrate levels may be modified to a more
 CC desirable level using the present invention. The present sequence is the
 CC amino acid sequence of a rice protein of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 476 AA;
 SQ Query Match 40.0%; Score 6; DB 7; Length 476;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 LRAYGF 12
 Db 449 LRAYGF 454
 RESULT 15
 ABO63378
 ID ABO63378 standard; protein; 484 AA.
 XX ABO63378;
 AC ABO63378;
 XX 29-JUL-2004 (first entry)
 DT Klebsiella pneumoniae polypeptide seqid 9895.
 XX Recombinant expression vector; transcription regulatory element;
 KW Klebsiella pneumoniae protein; antibacterial; vaccine.
 XX Klebsiella pneumoniae.
 OS US6610836-B1.
 PN 26-AUG-2003.
 PD

```

XX 27-JAN-2000; 2000US-00489039.
XX
XX 29-JAN-1999; 99US-0117747P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Breton GL, Osborne M;
XX
XX WPI; 2003-895346/82.
XX
XX N-PSDB; ACH96929.
XX
XX
XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
XX preparing a vaccine composition against Klebsiella pneumoniae.
XX
XX Disclosure; SEQ ID NO 9895; 932pp; English.
XX
XX The invention describes a new isolated nucleic acid encoding a Klebsiella
XX pneumoniae polypeptide. Also described are: a recombinant expression
XX vector comprising the nucleic acid, operably linked to a transcription
XX regulatory element; and a cell comprising the recombinant expression
XX vector. The nucleic acid is useful for preparing a vaccine composition
XX against Klebsiella pneumoniae. This is the amino acid sequence of a
XX Klebsiella pneumoniae polypeptide of the invention
XX
XX Sequence 484 AA;
XX
XX Query Match 40.0%; Score 6; DB 7; Length 484;
XX Best Local Similarity 100.0%; Pred. No. 1.9e+02;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 10 YGFWRLL 15
XX |||||
XX Db 450 YGFWRLL 455
XX
XX RESULT 16
XX AAU02903
XX ID AAU02903 standard; protein; 515 AA.
XX
XX AC AAU02903;
XX
XX DT 12-SEP-2001 (first entry)
XX
XX XX
XX XX Angiotensin converting enzyme (ACEV) splice variant protein #3.
XX
XX Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
XX granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
XX platelet-derived endothelial cell growth factor; cardiovascular disease;
XX cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C;
XX vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
XX myocardial infarction; coronary arterial thrombosis; renal disease;
XX diabetic nephropathy; immune disease; immune disorder; sarcoidosis;
XX multiple sclerosis; immune complex nephritis; deep vein thrombosis;
XX nonarocidotic pulmonary granulomatous disease; endothelial abnormality;
XX vascular disorder; asbestosis.
XX
XX XX
XX OS Homo sapiens.
XX
XX WO200136632-A2.
XX
XX XX
XX PD 25-MAY-2001.
XX
XX PF 17-NOV-2000; 2000WO-IL000766.
XX
XX PR 17-NOV-1999; 99IL-00132978.
XX
XX PR 10-DEC-1999; 99IL-00133455.
XX
XX (COMP-) COMPUGEN LTD.
XX
XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;
XX WPI; 2001-336004/35.
XX

```


PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou F, Goodrich R, Drmanac RT;
 DR WPI; 2001-442253/47.
 DR N-PSDB; AAI59212.
 XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX Example 5; SEQ ID NO 3201; 10078pp; English.
 PS The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AM38642-AM42213) with neurotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX
 SQ Sequence 615 AA;

Query Match 40.0%; Score 6; DB 4; Length 615;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SYQHD 7
 |||||
 Db 457 SYQHD 462

RESULT 18
 ADA50545
 ID ADA50545 standard; protein; 615 AA.
 XX
 AC ADA50545;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human factor XII.
 XX
 KW Coagulation protein; Factor XII; bypass activity; haemophilia A;
 KW haemophilia B; Factor XIIa; haemostatic.
 XX
 OS Homo sapiens.
 XX
 PN US2003073652-A1.
 XX
 PD 17-APR-2003.
 XX
 PF 17-MAY-2001; 2001US-00858909.
 XX
 PR 17-MAY-2000; 2000US-0205014P.
 XX
 PA (POLL/) POLLARD H.
 PA (POLL/) POLLARD B.
 XX
 PI Pollard H, Pollard B;
 XX
 DR WPI; 2003-644802/61.
 DR N-PSDB; ADA50544.
 XX
 PT Treating a subject with a coagulation protein, e.g. Factor XII
 PT polypeptide, comprises administering a gene for Factor XII or a cell
 PT transfected with full-length or a truncated version of Factor XII.

XX Claim 3; Page 9-10; 16pp; English.
 XX The present invention relates to a method for treating a subject with a
 CC gene encoding the coagulation protein, Factor XII. Bypass activity for
 CC haemophilia A and B can be generated by natural or recombinant Factor
 CC XIIa. A continuous source of Factor XIIa, provided by gene therapy, is
 CC therapeutic for haemophilia A and B. A gene encoding a Factor XII
 CC polypeptide can be administered to a patient by intravenous,
 CC intraarterial, intraportal, intracranial, intrapleural or intraperitoneal
 CC routes. Administering a gene encoding a Factor XII polypeptide is used
 CC for providing a subject with a coagulation protein and is useful for
 CC treating patients with haemophilia A and B. The present sequence
 CC represents human Factor XII.
 XX

SQ Sequence 615 AA;

Query Match 40.0%; Score 6; DB 7; Length 615;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SYQHD 7
 |||||
 Db 457 SYQHD 462

RESULT 19
 ADN04180
 ID ADN04180 standard; protein; 615 AA.
 XX
 AC ADN04180;

XX 01-JUL-2004 (first entry)
 DT
 XX Antipsoriatic protein sequence #285.
 DE
 XX antipsoriatic; gene therapy; psoriasis; diagnosis.
 KW
 OS Homo sapiens.
 XX
 PN WO2004028479-A2.
 XX
 PD 08-APR-2004.
 XX
 PF 25-SEP-2003; 2003WO-US030907.
 XX
 PR 25-SEP-2002; 2002US-0414006P.
 XX
 PA (GETH) GENENTECH INC.

XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
 PI Wu TD;
 XX
 DR WPI; 2004-305105/28.
 DR N-PSDB; ADN04179.
 XX
 PT New PRO nucleic acid or polypeptide, useful for preparing a
 PT pharmaceutical composition for diagnosing or treating psoriasis in a
 PT mammal.
 XX
 PS Claim 9; SEQ ID NO 574; 3069pp; English.
 XX The invention relates to novel polynucleotide and polypeptides for
 CC treating psoriasis or a sequence having at least 80% identity to the
 CC above sequences. The nucleic acid is useful for preparing a composition
 CC for diagnosing or treating psoriasis in a mammal. This sequence
 CC corresponds to one of the polypeptides of the invention.
 XX

SQ Sequence 615 AA;

Query Match 40.0%; Score 6; DB 8; Length 615;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SYQHDL 7
Db 457 SYQHDL 462

RESULT 20
ADM98382
ID ADM98382 standard; protein; 615 AA.
XX
XX ADM98382;
XX
XX 15-JUL-2004 (first entry)
XX Human factor XII protein.
XX
XX haenostatic; gene therapy; coagulation protein; fluid communication;
KW Factor VII; bloodstream; Factor XIIa; bleeding; coagulation agent;
KW haemophilia A; haemophilia B; immunosuppression; human.
XX
XX Homo sapiens.
OS
PN US2004077538-A1.
XX
XX 22-APR-2004.
PD
XX 02-JUN-2003; 2003US-00449132.
PF
XX 17-MAY-2000; 2000US-0205014P.
PR
XX 17-MAY-2001; 2001US-00858909.
XX
XX (POLL/) POLLARD H.
XX (POLL/) POLLARD B.
XX
XX Pollard H, Pollard B;
XX
XX WPI; 2004-374247/35.
DR N-PSDB; ADM98381.
XX
XX Treating hemophilia in subject with Factor XIIa coagulation protein, by
PT providing Factor XIIa gene as construct, transfecting host cells with
PT construct, adding cells expressing protein to porous chamber, implanting
PT chamber into subject.
XX
XX Claim 3; SEQ ID NO 2; 18pp; English.

XX The invention describes a method of treating (M1) a subject with a
CC coagulation protein. The method involves providing a gene for a
CC coagulation protein prepared as an expression vector construct,
CC transfecting host cells with the construct, adding the host cells
CC expressing the coagulation protein to a porous chamber, and implanting
CC the chamber into the subject, where the chamber is implanted in fluid
CC communication with the bloodstream and inactive Factor VII in the
CC bloodstream passing through the chamber becomes activated upon contact
CC with the coagulation protein or the method optionally involves providing
CC a gene prepared as a naked DNA vector and administering the vector to the
CC subject in a therapeutically effective amount, where the gene encodes a
CC Factor XIIa polypeptide. Also described are: a Factor XIIa-coated solid
CC support (1) comprising a recombinant Factor XIIa polypeptide; a Factor
CC XIIa-expressing cell (11), comprising an expression vector construct
CC containing a gene encoding Factor XII polypeptide; treating or preventing
CC (M2) bleeding involving administering to a subject suffering from
CC bleeding, a coagulation agent, where the coagulation agent comprises a
CC therapeutically effective amount of a Factor XIIa polypeptide, and where
CC the coagulation agent is administered in an amount effective to treat or
CC prevent bleeding; and a kit (111), comprising an applicator and a
CC coagulation agent, where the coagulation agent comprises Factor XIIa
CC polypeptide. (M1) or (M2) are effective in treating a subject suffering
CC from haemophilia with Factor XIIa coagulation protein. The method is
CC effective in treating both haemophilia A and haemophilia B. Antibodies
CC against Factor XIIa or Factor VII are not produced and hence
CC immunosuppression during the treatment is not necessary. This is the
CC amino acid sequence of human factor XII.

XX SQ Sequence 615 AA;
Query Match 40.0%; Score 6; DB 8; Length 615;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SYQHDL 7
Db 457 SYQHDL 462

RESULT 21
ABM80709
ID ABM80709 standard; protein; 615 AA.
XX
XX AC ABM80709;
XX
XX 18-NOV-2004 (first entry)
XX
XX Tumour-associated antigenic target (TAT) polypeptide PRO36372, SEQ.1829.
XX
XX Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic.
XX
XX Homo sapiens.
OS
XX WO2004030615-A2.
PN
XX 15-APR-2004.
PD
XX 29-SEP-2003; 2003WO-US028547.
PF
XX 02-OCT-2002; 2002US-0414971P.
PR
XX (GETH) GENENTECH INC.
XX
XX Wu TD, Zhang Z, Zhou Y;
XX
XX WPI; 2004-347921/32.
DR N-PSDB; ACN38363.
XX
XX New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
XX Claim 12; SEQ ID NO 1829; 7273pp; English.

XX The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence

CC represents a TAT polypeptide of the invention
SQ Sequence 615 AA;
Query Match 40.0%; Score 6; DB 8; Length 615;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 SYQHDL 7
DB 457 SYQHDL 462
RESULT 22
AAU11271
ID AAU11271 standard; protein; 621 AA.
XX AC AAU11271;
XX DT 12-MAR-2002 (first entry)
XX DE Human coagulation factor XII (F12) variant polypeptide.
XX KW Human; coagulation factor XII; F12; haplotyping; haplotype pair;
KW single nucleotide polymorphism; genotyping; gene therapy; drug screening;
KW coronary artery disease; liver disease; spontaneous abortion; cardiant;
KW Alzheimer's disease; blood coagulation; hepatotropic; neuroprotective;
KW nootropic; coagulant; antiabortive.
XX OS Homo sapiens.
FH Key Location/Qualifiers
FT Misc-difference 140
FT /note= "Leu substituted by Val"
FT Misc-difference 343
FT /note= "Ala substituted by Pro"
FT Misc-difference 448
FT /note= "Arg substituted by Cys"
XX FT
XX PN WO200179228-A2.
XX PD 25-OCT-2001.
XX PF 13-APR-2001; 2001WO-US012257.
XX PR 14-APR-2000; 2000US-0197837P.
XX PA (GENA-) GENAISSANCE PHARM INC.
XX PI Bentivegna SC, Chew A, Choi JY, Nandabalan K;
XX WPI; 2002-075061/10.
XX PT Novel isolated human coagulation factor XII polynucleotide, F12 useful
PT for treatment of e.g. coronary artery disease, comprises a sequence which
PT is a polymorphic variant of a reference sequence for F12 gene or its
PT fragment.
XX PS Disclosure; Page; 72pp; English.
XX CC The invention relates to single nucleotide polymorphisms in the gene
CC encoding the human coagulation factor XII (F12) polypeptide. A method for
CC haplotyping the F12 gene in an individual comprises identifying the
CC nucleotide at one or more polymorphic sites and determining whether one
CC of the copies of the gene is defined by one of the F12 haplotypes given
CC in the specification or whether both copies are defined by a haplotype
CC pair. This method is useful in genotyping, whereby all possible haplotype
CC pairs can be assigned to specific genotypes. An association between a
CC trait and a haplotype or haplotype pair of the F12 gene can be identified
CC by comparing the frequency of the haplotype or haplotype pair in a
CC population exhibiting the trait with the frequency of the haplotype or
CC haplotype pair in a reference population, where a higher haplotype
CC frequency in the trait population indicates the trait is associated with

CC the haplotype or haplotype pair. F12 and its corresponding DNA are used
CC for studying the expression and function of F12, for use in screening for
CC candidate drugs to treat disorders related to F12 activity such as
CC coronary artery disease, liver disease, spontaneous abortion, Alzheimer's
CC disease and other diseases associated with defects in blood coagulation.
CC The sequences are also useful for studying the effect of variation on the
CC biological activity of F12 as well as on the binding affinity of
CC candidate drugs targeting F12. This sequence represents a human F12
CC variant polypeptide
XX SQ Sequence 621 AA;
Query Match 40.0%; Score 6; DB 5; Length 621;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 SYQHDL 7
DB 457 SYQHDL 462
RESULT 23
AAU11270
ID AAU11270 standard; protein; 621 AA.
XX AC AAU11270;
XX DT 12-MAR-2002 (first entry)
XX DE Human coagulation factor XII (F12) polypeptide.
XX KW Human; coagulation factor XII; F12; haplotyping; haplotype pair;
KW single nucleotide polymorphism; genotyping; gene therapy; drug screening;
KW coronary artery disease; liver disease; spontaneous abortion; cardiant;
KW Alzheimer's disease; blood coagulation; hepatotropic; neuroprotective;
KW nootropic; coagulant; antiabortive.
XX OS Homo sapiens.
XX PN WO200179228-A2.
XX PD 25-OCT-2001.
XX PF 13-APR-2001; 2001WO-US012257.
XX PR 14-APR-2000; 2000US-0197837P.
XX PA (GENA-) GENAISSANCE PHARM INC.
XX PI Bentivegna SC, Chew A, Choi JY, Nandabalan K;
XX WPI; 2002-075061/10.
XX DR N-PSDB; AAS16944.
XX PT Novel isolated human coagulation factor XII polynucleotide, F12 useful
PT for treatment of e.g. coronary artery disease, comprises a sequence which
PT is a polymorphic variant of a reference sequence for F12 gene or its
PT fragment.
XX PS Claim 29; Fig 3; 72pp; English.
XX CC The invention relates to single nucleotide polymorphisms in the gene
CC encoding the human coagulation factor XII (F12) polypeptide. A method for
CC haplotyping the F12 gene in an individual comprises identifying the
CC nucleotide at one or more polymorphic sites and determining whether one
CC of the copies of the gene is defined by one of the F12 haplotypes given
CC in the specification or whether both copies are defined by a haplotype
CC pair. This method is useful in genotyping, whereby all possible haplotype
CC pairs can be assigned to specific genotypes. An association between a
CC trait and a haplotype or haplotype pair of the F12 gene can be identified
CC by comparing the frequency of the haplotype or haplotype pair in a
CC population exhibiting the trait with the frequency of the haplotype or
CC haplotype pair in a reference population, where a higher haplotype
CC frequency in the trait population indicates the trait is associated with

CC frequency in the trait population indicates the trait is associated with
 CC the haplotype or haplotype pair. F12 and its corresponding DNA are used
 CC for studying the expression and function of F12, for use in screening for
 CC candidate drugs to treat disorders related to F12 activity such as
 CC coronary artery disease, liver disease, spontaneous abortion, Alzheimer's
 CC disease and other diseases associated with defects in blood coagulation.
 CC The sequences are also useful for studying the effect of variation on the
 CC biological activity of F12 as well as on the binding affinity of
 CC candidate drugs targeting F12. This sequence represents the human F12
 CC polypeptide
 XX
 XX Sequence 621 AA;

SQ Sequence 621 AA;
 Query Match 40.0%; Score 6; DB 5; Length 621;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SYQHDL 7
 DB 457 SYQHDL 462

RESULT 24
 ADI32366
 ID ADI32366 standard; protein; 631 AA.

AC ADI32366;
 XX
 XX 22-APR-2004 (first entry)
 XX
 XX Rice phosphoribosylaminoimidazole carboxylase protein.
 DE
 DE Phosphoribosylaminoimidazole carboxylase; genetic mapping;
 KW physical mapping; AIR; rice; plant; enzyme.
 XX

OS Oryza sativa.

XX US2003229916-A1.

XX 11-DEC-2003.

XX 19-DEC-2002; 2002US-00323167.

XX 19-DEC-2001; 2001US-0341955P.

XX (ALIE/) ALLEN S M.

XX (OROZ/) OROZCO E M.

XX (LEEJ/) LEE J.

XX Allen SM, Orozco EM, Lee J;

XX WPI; 2004-052053/05.

XX N-PSDB; ADI32365.

XX New polynucleotide comprising a sequence encoding a polypeptide having

PT phosphoribosylaminoimidazole carboxylase activity, useful as a probe for

PT genetic or physical mapping.

XX Claim 14; SEQ ID NO 8; 44pp; English.

XX The invention relates to a nucleic acid fragment encoding

CC phosphoribosylaminoimidazole (AIR) carboxylase enzyme. The nucleic acid

CC molecule of the invention is useful as a probe for genetic or physical

CC mapping. The present sequence is rice phosphoribosylaminoimidazole

CC carboxylase protein.
 XX
 XX Sequence 631 AA;

Query Match 40.0%; Score 6; DB 8; Length 631;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 YQHDLR 8

Db 600 YQHDLR 605

RESULT 25
 ABB82636
 ID ABB82636 standard; protein; 653 AA.

XX ABB82636;

XX 19-FEB-2003 (first entry)

XX PERV envelope protein encoding DNA.

XX Porcine endogenous retrovirus; PERV-A; virucide; infection; viral; PERV;
 KW envelope protein.

XX Porcine endogenous retrovirus.

XX WO200286060-A2.

XX 31-OCT-2002.

XX 19-APR-2002; 2002WO-US012085.

XX 20-APR-2001; 2001US-0285103P.

XX 21-DEC-2001; 2001US-00029656.

XX (IMME-) IMMERGE BIOTHERAPEUTICS INC.

XX Patience C, Ericsson T, Oldmixon BA;

XX WPI; 2003-067636/06.

XX N-PSDB; ABV75054.

XX New porcine endogenous retrovirus (PERV)-A receptor polypeptides and
 PT polynucleotides, useful for screening agents that block viral infection
 PT (which are useful for inhibiting PERV infection) or detecting tissues
 PT infected with PERV.

XX Disclosure; Page 57-59; 71pp; English.

XX The invention relates to Porcine endogenous retrovirus (PERV)-A receptor
 CC polypeptides found on primate cells including human and baboon and
 CC encoding polynucleotides. The polypeptides are useful in both in vitro or
 CC in vivo screening assays and models for agents that block viral
 CC infection. These are also useful for detecting tissues infected with
 CC PERV. Anti-PERV antibodies are useful for inhibiting PERV infection in
 CC tissue, particularly in tissues for transplantation. The present sequence
 CC represents the PERV-A envelope protein
 XX Sequence 653 AA;

Query Match 40.0%; Score 6; DB 6; Length 653;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRAYGF 12

DB 108 LRAYGF 113

RESULT 26
 ABB82635
 ID ABB82635 standard; protein; 653 AA.

XX ABB82635;

XX 19-FEB-2003 (first entry)

XX PERV envelope protein.

XX Porcine endogenous retrovirus; PERV-A; virucide; infection; viral; PERV;

```

KW envelope protein.
XX
OS Porcine endogenous retrovirus.
XX
XX WO200286060-A2.
XX
XX 31-OCT-2002.
XX
XX 19-APR-2002; 2002WO-US012085.
XX
XX 20-APR-2001; 2001US-0285103P.
XX
XX 21-DEC-2001; 2001US-00029656.
XX
XX (IMME-) IMMERGE BIOTHERAPEUTICS INC.
XX
XX Patience C, Ericsson T, Oldmixon BA;
XX
XX WPI; 2003-067636/06.
XX
XX N-PSDB; ABV75053.
XX
XX New porcine endogenous retrovirus (PERV)-A receptor polypeptides and
XX PT polynucleotides, useful for screening agents that block viral infection
XX PT (which are useful for inhibiting PERV infection) or detecting tissues
XX PT infected with PERV.
XX
XX Disclosure; Page 54-56; 71pp; English.
XX
XX The invention relates to Porcine endogenous retrovirus (PERV)-A receptor
XX CC polypeptides found on primate cells including human and baboon and
XX CC encoding polynucleotides. The polypeptides are useful in both in vitro or
XX CC in vivo screening assays and models for agents that block viral
XX CC infection. These are also useful for detecting tissues infected with
XX CC PERV. Anti-PERV antibodies are useful for inhibiting PERV infection in
XX CC tissue, particularly in tissues for transplantation. The present sequence
XX CC represents the PERV-A envelope protein
XX
XX Sequence 653 AA;
SQ
Query Match 40.0%; Score 6; DB 6; Length 653;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7 LRAYGF 12
DB 108 LRAYGF 113
|||||
|||||

RESULT 27
AAW32095
ID AAW32095 standard; protein; 657 AA.
XX
XX AAW32095;
XX
XX 27-AUG-2003 (revised)
XX
XX 09-FEB-1998 (first entry)
XX
XX Porcine retrovirus partial ENV protein from defective protein.
XX
XX Retrovirus; porcine; ENV protein; infectious; provirus;
XX KW xenotransplantation; organ transplant; donor; activated virus; PCR.
XX
XX Pig endogenous retrovirus.
XX
XX Key Location/Qualifiers
XX FT Protein 1..657
XX FT /label= ENV_protein
XX FT /note= "fragment containing putative partial ENV viral
XX PN protein"
XX
XX WO9721836-A1.
XX
XX 19-JUN-1997.
XX

```

```

PF 13-DEC-1996; 96WO-US019680.
XX
XX 14-DEC-1995; 95US-00572645.
XX
XX (GEO) GEN HOSPITAL CORP.
XX
XX Fishman JA;
XX
XX WPI; 1997-332804/30.
XX
XX N-PSDB; AAT74883.
XX
XX New nucleic acid from porcine retroviruses - used for detecting viruses
XX PT in transplant or other tissue and for assessing risk of transmitting
XX PT infection to graft recipient.
XX
XX Claim 16; Fig 2; 128pp; English.
XX
XX This sequence contains a putative partial viral ENV protein from a
XX CC defective swine retrovirus found in PK-15 cells. There are a few in frame
XX CC stop codons and apparent frame shifts in the given cDNA coding sequence
XX CC which alter features of the translation (see AAW32092-W32094 for other
XX CC parts of the sequence). This sequence and PCR fragments generated from
XX CC the sequence (see AAT74812-T74882) could be used to screen organs for
XX CC porcine retroviruses prior to xenotransplantation. Transplantation can
XX CC increase the likelihood of retroviral activation if intact and infectious
XX CC proviruses are present. The porcine retroviral sequence can be used to
XX CC generate probes to determine the level (e.g. copy number) of intact (i.e.
XX CC potentially replicating) porcine provirus sequences in a strain of
XX CC xenograft transplanting donors. It can be used to detect mutations,
XX CC genetic lesions or viral recombinants and also to determine the
XX CC histological localization of activated retrovirus. Using Polymerase Chain
XX CC Reaction DNA Quantitation (PQ) on blood mononuclear cells, infectivity
XX CC titration and susceptibility testing can be performed. Ultimately animal
XX CC donors without intact porcine retroviral sequences or a lower copy number
XX CC of viral elements could be selected. (Updated on 27-AUG-2003 to correct
XX CC OS field.)
XX
XX Sequence 657 AA;
SQ
Query Match 40.0%; Score 6; DB 2; Length 657;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7 LRAYGF 12
DB 105 LRAYGF 110
|||||
|||||

RESULT 28
ADS73445
ID ADS73445 standard; protein; 657 AA.
XX
XX ADS73445;
XX
XX 16-DEC-2004 (first entry)
XX
XX Swine retroviral env protein fragment.
XX
XX Swine retroviral protein; immunosuppressive; gene therapy; env protein.
XX KW
XX Porcine endogenous retrovirus.
XX
XX US2004185435-A1.
XX
XX 23-SEP-2004.
XX
XX 26-NOV-2003; 2003US-00723552.
XX
XX 14-DEC-1995; 95US-00572645.
XX
XX 13-DEC-1996; 96US-00766528.
XX
XX 14-SEP-2000; 2000US-00661858.
XX
XX (GEO) GEN HOSPITAL CORP.
XX

```

XX Fishman JA;
XX WPI; 2004-689179/57.
DR N-PSDB; ADS73368.
XX New porcine retroviral polypeptide encoded by a nucleic acid, useful in
PT evaluating an immunosuppressive treatment for the ability to activate a
PT retrovirus, such as an endogenous porcine retrovirus.
XX Claim 8; Fig 2; 83pp; English.
XX The present invention relates to the swine retroviral polypeptides and
CC their encoding nucleic acids. The methods and compositions of the present
CC invention are useful for screening a cell or tissue, e.g. a heart, lung,
CC liver, bone marrow, kidney, brain cells, neural tissue, pancreas and
CC intestinal tissue xenograft, for the presence or expression of a swine or
CC miniature swine retrovirus or retroviral sequence. The invention is also
CC useful in evaluating an immunosuppressive treatment for the ability to
CC activate a retrovirus such as an endogenous porcine retrovirus. The
CC invention is also useful in gene therapy. The present sequence is the
CC swine retroviral env protein fragment.
XX Sequence 657 AA;
SQ Query Match 40.0%; Score 6; DB 8; Length 657;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 LRAYGF 12
DB 105 LRAYGF 110
RESULT 29
ID AAB35112 standard; protein; 659 AA.
AC AAB35112;
XX 27-MAR-2001 (first entry)
DT PERV-1-15 env protein SEQ ID NO: 3.
DE Xenotransplantation; infectious agent; vaccine.
XX Porcine endogenous retrovirus.
OS
XX WO200071726-A1.
XX 30-NOV-2000.
XX 24-MAY-2000; 2000WO-US014296.
XX 24-MAY-1999; 99US-0135631P.
XX (MAYO-) MAYO MEDICAL VENTURES.
XX Federspiel MJ;
XX WPI; 2001-032041/04.
XX Inhibiting or preventing infectious agent transmission in mammalian
PT transplant recipients, by introducing recombinant DNA comprising DNA
PT encoding extracellular proteins of the agent into donor cells, such as
PT swine cells.
XX Disclosure; Page 91-93; 144pp; English.
XX The present invention provides a method to prevent the transmission of
CC infectious agents during xenotransplantation. This involves introducing
CC to donor swine cells a recombinant DNA encoding a peptide fragment from
CC the infectious agent, and then introducing these cells into the

CC transplant recipient
XX Sequence 659 AA;
SQ Query Match 40.0%; Score 6; DB 4; Length 659;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 LRAYGF 12
DB 108 LRAYGF 113
RESULT 30
ID AAW85453 standard; protein; 660 AA.
XX AAW85453;
AC AAW85453;
XX 25-FEB-1999 (first entry)
DT Pig endogenous retrovirus (PERV)-A envelope (env) protein.
DE Pig endogenous retrovirus (PERV)-A envelope (env) protein.
XX Pig endogenous retrovirus; PERV-A; envelope protein; PERV-B; subtype;
KW probe; primer; detection; retrovirus; human tissue; xenotransplant;
XX primary porcine tissue; human cell line; porcine cell line.
OS Pig endogenous retrovirus.
XX WO9853104-A2.
XX 26-NOV-1998.
XX 18-MAY-1998; 98WO-GB001428.
XX 16-MAY-1997; 97GB-00010154.
XX (MEDI-) MEDICAL RES COUNCIL.
XX Stoye JP, Weiss RA;
XX WPI; 1999-045324/04.
XX N-PSDB; AAV82748.
XX Newly isolated nucleic acid probe capable of hybridising to either the
PT PERV-A or PERV-B env gene - useful in the detection of retroviruses, and
PT their subtypes, in a sample of porcine/human tissue.
XX Example 1; Page 21; 36pp; English.
XX The present sequence represents a Pig endogenous retrovirus (PERV)-A
CC envelope protein. PERV exists in two different subtypes, PERV-A and PERV-
CC B. The differences are reflected in sequence divergence in the envelope
CC genes. Probes and primers can be derived from the envelope (env) genes of
CC PERV-A and PERV-B. The probes and primers are used in a method to detect
CC retroviruses in a sample of porcine/human tissue, particularly primary
CC porcine tissue and human cell lines that have been cultivated in the
CC presence of a porcine cell line, or human tissue from a patient with a
CC xenotransplant. Subtype of PERV in a sample containing one of the PERV
CC env genes can also be determined
XX Sequence 660 AA;
SQ Query Match 40.0%; Score 6; DB 2; Length 660;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 LRAYGF 12
DB 108 LRAYGF 113
RESULT 31

```

AAB35113
ID AAB35113 standard; protein; 660 AA.
XX
XX AC AAB35113;
XX
XX DT 27-MAR-2001 (first entry)
XX
XX DE PERV-A env protein SEQ ID NO: 4.
XX
XX KW Xenotransplantation; infectious agent; vaccine.
XX
XX OS Porcine endogenous retrovirus.
XX
XX PN WO200071726-A1.
XX
XX PD 30-NOV-2000.
XX
XX PF 24-MAY-2000; 2000WO-US014296.
XX
XX PR 24-MAY-1999; 99US-0135631P.
XX
XX PA (MAYO-) MAYO MEDICAL VENTURES.
XX
XX PI Federspiel MJ;
XX
XX DR WPI; 2001-032041/04.
XX
XX PT Inhibiting or preventing infectious agent transmission in mammalian
PT transplant recipients, by introducing recombinant DNA comprising DNA
PT encoding extracellular proteins of the agent into donor cells, such as
PT swine cells.
XX
XX PS Disclosure; Page 93-95; 144pp; English.
XX
XX CC The present invention provides a method to prevent the transmission of
CC infectious agents during xenotransplantation. This involves introducing
CC to donor swine cells a recombinant DNA encoding a peptide fragment from
CC the infectious agent, and then introducing these cells into the
CC transplant recipient
XX
XX SQ Sequence 660 AA;

Query Match 40.0%; Score 6; DB 4; Length 660;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LRAYGF 12
Db 108 LRAYGF 113

RESULT 32
AAB70632
ID AAB70632 standard; protein; 660 AA.
XX
XX AC AAB70632;
XX
XX DT 15-MAY-2001 (first entry)
XX
XX DE Porcine endogenous retrovirus envelope protein PERVA.
XX
XX KW Porcine endogenous retrovirus; PoERV; gag; env; detection; antigenic;
KW immunogenic; antiviral; vaccine; antiserum; viral infection.
XX
XX OS Porcine endogenous retrovirus.
XX
XX PN WO200112816-A2.
XX
XX PD 22-FEB-2001.
XX
XX PF 16-AUG-2000; 2000WO-GB003159.
XX
XX PR 18-AUG-1999; 99GB-00019604.

Query Match 40.0%; Score 6; DB 4; Length 660;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LRAYGF 12
Db 108 LRAYGF 113

RESULT 33
AAB73284
ID AAB73284 standard; protein; 1069 AA.
XX
XX AC AAB73284;
XX
XX DT 23-MAY-2001 (first entry)
XX
XX DE Defective retroviral genome protein #3 isolated from PK-15 cell line.
XX
XX KW Retrovirus; graft transplantation; xenotransplantation; PK-15 cell line.
XX
XX OS Unidentified.
XX
XX PN US6190861-B1.
XX
XX PD 20-FEB-2001.
XX
XX PF 13-DEC-1996; 96US-00766528.
XX
XX PR 14-DEC-1995; 95US-00572645.
XX
XX PA (GEHO ) GEN HOSPITAL CORP.
XX
XX PI Fishman JA;
XX
XX DR WPI; 2001-256211/26.
XX
XX DR N-PSDB; AAF77726.
XX
XX PT Assessing risk of endogenous retroviruses in clinical practice and in
PT xenotransplantation, comprises using probe sequences derived from swine
PT or miniature swine retroviral genome.
XX

```

```

XX
XX PA (QUIP-) QUIP TECHNOLOGY LTD.
XX
XX PI Galbraith D, Kelly H, Smith K;
XX
XX DR WPI; 2001-211223/21.
XX
XX PT New porcine endogenous retrovirus polypeptide fragment with retroviral
PT specific antigenic or immunogenic activity, for detection of retroviral
PT antibodies in a sample and in therapy or diagnosis.
XX
XX PS Claim 7; Fig 2; 43pp; English.
XX
XX CC The present invention describes a porcine endogenous retrovirus (PoERV)
CC polypeptide fragment (I) which has PoERV specific antigenic or
CC immunogenic activity. Also described are: (1) an antiserum (II) specific
CC to (I); (2) a PoERV specific antibody (III) or its fragment raised
CC against (I); (3) use of (I) in detecting (III) in a sample; (4) use of
CC (III) in detecting PoERV in a sample; (5) use of (I) or (III) in therapy
CC or diagnosis; and (6) an assay kit (IV) comprising (I) or (III) for use
CC in detection of PoERV in a sample. (I) has antiviral activity, and can be
CC used in vaccine production. (I) is useful in the detection of PoERV
CC antibodies (III) in a sample and (III) is useful in the detection of
CC antiserum (II) to (I) and kits (IV) comprising (I) or (III) are useful in
CC detection of specific PoERV types. Viral infection can be monitored and
CC by using (I) or (III). The present sequence represents a specifically
CC claimed PoERV envelope (env) protein sequence from the present invention
XX
XX SQ Sequence 660 AA;

Query Match 40.0%; Score 6; DB 4; Length 660;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LRAYGF 12
Db 108 LRAYGF 113

RESULT 33
AAB73284
ID AAB73284 standard; protein; 1069 AA.
XX
XX AC AAB73284;
XX
XX DT 23-MAY-2001 (first entry)
XX
XX DE Defective retroviral genome protein #3 isolated from PK-15 cell line.
XX
XX KW Retrovirus; graft transplantation; xenotransplantation; PK-15 cell line.
XX
XX OS Unidentified.
XX
XX PN US6190861-B1.
XX
XX PD 20-FEB-2001.
XX
XX PF 13-DEC-1996; 96US-00766528.
XX
XX PR 14-DEC-1995; 95US-00572645.
XX
XX PA (GEHO ) GEN HOSPITAL CORP.
XX
XX PI Fishman JA;
XX
XX DR WPI; 2001-256211/26.
XX
XX DR N-PSDB; AAF77726.
XX
XX PT Assessing risk of endogenous retroviruses in clinical practice and in
PT xenotransplantation, comprises using probe sequences derived from swine
PT or miniature swine retroviral genome.
XX

```

PS Disclosure; Fig 2; 127pp; English.

XX The present invention relates to a method for screening a cell or tissue

CC for the presence or expression of a retrovirus (RV), comprising

CC contacting a target nucleic acid from the cell or tissue with a second

CC nucleic acid from the present invention (e.g. AAF77726 or a fragment

CC thereof). The method is useful for RV detection and to assess graft

CC transplantation risk. Screening of animals allows the elimination of

CC donors with active replication of known viruses. Inactive proviruses can

CC be detected and inactivated, allowing identification and elimination of

CC potential human pathogens derived from swine in a manner not possible in

CC the outbred human organ donor population and is important to the

CC development of human xenotransplantation

XX

SQ Sequence 1069 AA;

Query Match 40.0%; Score 6; DB 4; Length 1069;

Best Local Similarity 100.0%; Pred. No. 3.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRAYGF 12

DB 517 LRAYGF 522

RESULT 34

ADN73089

ID ADN73089 standard; protein; 1084 AA.

XX

AC ADN73089;

XX

DT 15-JUL-2004 (first entry)

XX

DE Thale cress protein upregulated in E2Fa/Dpa expressing plants SeqID 984.

XX

KW plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;

KW animal feed product; thale cress; cell wall biosynthesis;

KW nitrogen metabolism; carbon metabolism.

XX

OS Arabidopsis thaliana.

XX

PN WO2004035798-A2.

XX

PD 29-APR-2004.

XX

PF 20-OCT-2003; 2003WO-EP011658.

XX

PR 18-OCT-2002; 2002EP-00079408.

XX

PA (CROP-) CROPDISEIGN NV.

XX

PI Inze D, De Veylder L, Vlieghe K;

XX

DR WPI; 2004-348466/32.

DR N-PSDB; ADN73088.

XX

PT Altering plant characteristics, useful for producing plants for enzyme or

PT pharmaceutical production comprising modifying in a plant, expression of

PT one or more nucleic acids and/or modifying level or activity of one or

PT more proteins.

XX

PS Claim 1; SEQ ID NO 984; 134pp; English.

XX

CC This invention relates to a novel method for altering one or more plant

CC characteristics. Specifically, it refers to identifying genes that are up

CC - or down-regulated in transgenic plants overexpressing the heterodimeric

CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to

CC alter plant characteristics accordingly. The present invention describes

CC generating transgenic plants for the production of growth regulators,

CC enzymes, therapeutics, pharmaceuticals and animal feed products, where

CC the altered plant characteristics are selected from increased yield or

CC biomass, enhanced survival capacity, stress tolerance, plant architecture

CC or physiology, altered endoreplication, biochemistry, signal

CC

transduction, storage lipid mobilisation and/or altered photosynthesis,

CC each relative to the corresponding wild type plants. Accordingly, these

CC sequences can also be useful as positive or negative selectable markers

CC during transformation of cells or tissues. The identified genes play a

CC role in a variety of biological processes such as DNA replication, cell

CC wall biosynthesis, nitrogen and/ or carbon metabolism or they function as

CC transcription factors. This polypeptide sequence is thale cress protein

CC expressed by a gene upregulated 1.3 fold or more in plants overexpressing

CC the E2Fa/Dpa transcription factor, given in an exemplification of the

CC invention.

SQ Sequence 1084 AA;

Query Match 40.0%; Score 6; DB 8; Length 1084;

Best Local Similarity 100.0%; Pred. No. 3.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 DLRAYG 11

DB 195 DLRAYG 200

RESULT 35

ADR86081

ID ADR86081 standard; protein; 1420 AA.

XX

AC ADR86081;

XX

DT 04-NOV-2004 (first entry)

XX

DE Aspergillus fumigatus essential gene protein #131.

XX

KW Fungicide; Aspergillus fumigatus infection; Farmer's lung disease;

KW drug screening.

XX

OS Aspergillus fumigatus.

XX

PN WO2004067709-A2.

XX

PD 12-AUG-2004.

XX

PF 16-JAN-2004; 2004WO-US001099.

XX

PR 17-JAN-2003; 2003US-0441281P.

XX

PR 13-JUN-2003; 2003US-0478196P.

XX

PA (ELIT-) ELITRA PHARM INC.

PA (ELIT-) ELITRA CANADA LTD.

XX

PI Jiang B, Hu W, Lemieux S, Roemer T;

XX

DR WPI; 2004-594200/57.

DR N-PSDB; ADR85494.

XX

PT New purified or isolated Aspergillus fumigatus nucleic acid molecule

PT encoding a gene product, useful for diagnosing and/or treating invasive

PT fungal infections, such as Farmer's lung disease.

XX

PS Claim 1; SEQ ID NO 3131; 164pp; English.

XX

CC The present invention relates to Aspergillus fumigatus genes that are

CC essential and are potential targets for drug screening. The methods and

CC compositions of the present invention are useful for diagnosing and/or

CC treating invasive Aspergillus fumigatus infection, including the allergic

CC forms of the disease, such as Farmer's lung disease. They can also be

CC used in various drug discovery purposes, such as expression of the

CC recombinant protein, hybridization assay and construction of nucleic acid

CC arrays. The present sequence represents an Aspergillus fumigatus

CC essential gene protein sequence, used during diagnosis and drug

CC development in the invention. These genes share a high degree of sequence

CC conservation with known essential genes of candida albicans. The sequence

CC data for this patent is not represented in the printed specification, but

CC was obtained in electronic format from WIPO.

XX
SQ Sequence 1420 AA;
Query Match 40.0%; Score 6; DB 8; Length 1420;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 RAYGFW 13
Db 1053 RAYGFW 1058
|||||

Search completed: March 18, 2005, 18:46:26
Job time : 167 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 18:30:52 ; Search time 137 Seconds
(without alignments)
36.191 Million cell updates/sec

Title: US-09-623-038-6

Perfect score: 87

Sequence: 1 RSYQHDLRAYGFWRL 15

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1401741 seqs, 330541175 residues

Total number of hits satisfying chosen parameters: 1401741

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|-------------------|
| 1 | 49 | 56.3 | 554 | US-10-437-963-169736 | Sequence 169736, |
| 2 | 44 | 50.6 | 1356 | US-10-197-824-35 | Sequence 35, Appl |
| 3 | 44 | 50.6 | 1593 | US-10-197-824-19 | Sequence 19, Appl |
| 4 | 43 | 49.4 | 256 | US-10-156-761-9242 | Sequence 9242, Ap |
| 5 | 43 | 49.4 | 465 | US-09-957-005-11 | Sequence 11, Appl |
| 6 | 43 | 49.4 | 466 | US-10-282-122A-59950 | Sequence 59950, A |
| 7 | 43 | 49.4 | 530 | US-09-966-997-13 | Sequence 13, Appl |
| 8 | 42.5 | 48.9 | 413 | US-10-425-114-47984 | Sequence 59634, A |
| 9 | 42 | 48.3 | 370 | US-10-437-963-194505 | Sequence 194505, |
| 10 | 42 | 48.3 | 558 | US-10-767-701-42825 | Sequence 42825, A |
| 11 | 42 | 48.3 | 891 | US-10-437-963-121169 | Sequence 121169, |
| 12 | 41 | 47.1 | 129 | US-10-424-599-211706 | Sequence 211706, |
| 13 | 41 | 47.1 | 497 | US-10-293-958-5 | Sequence 5, Appl |

| | | | | | |
|----|------|------|------|----------------------|-------------------|
| 14 | 41 | 47.1 | 544 | US-10-106-275-6 | Sequence 6, Appli |
| 15 | 41 | 47.1 | 573 | US-10-437-963-129380 | Sequence 129380, |
| 16 | 41 | 47.1 | 692 | US-10-389-566-1468 | Sequence 1468, Ap |
| 17 | 41 | 47.1 | 817 | US-10-389-566-1481 | Sequence 1481, Ap |
| 18 | 41 | 47.1 | 1479 | US-10-437-963-181030 | Sequence 181030, |
| 19 | 40 | 46.0 | 76 | US-10-424-599-281147 | Sequence 281147, |
| 20 | 40 | 46.0 | 150 | US-10-156-761-12936 | Sequence 12936, A |
| 21 | 40 | 46.0 | 151 | US-09-765-815-13 | Sequence 13, Appl |
| 22 | 40 | 46.0 | 223 | US-10-380-132-18 | Sequence 18, Appl |
| 23 | 40 | 46.0 | 334 | US-09-287-599-2 | Sequence 2, Appli |
| 24 | 40 | 46.0 | 478 | US-10-425-114-47984 | Sequence 47984, A |
| 25 | 40 | 46.0 | 490 | US-10-424-599-166245 | Sequence 166245, |
| 26 | 40 | 46.0 | 618 | US-10-235-403-166 | Sequence 166, App |
| 27 | 40 | 46.0 | 619 | US-09-934-455-142 | Sequence 142, App |
| 28 | 40 | 46.0 | 619 | US-10-225-066A-894 | Sequence 894, App |
| 29 | 40 | 46.0 | 619 | US-10-374-780A-164 | Sequence 164, App |
| 30 | 40 | 46.0 | 619 | US-10-412-699B-566 | Sequence 566, App |
| 31 | 40 | 46.0 | 677 | US-10-437-963-183525 | Sequence 183525 |
| 32 | 40 | 46.0 | 920 | US-09-934-868-52 | Sequence 52, Appl |
| 33 | 39.5 | 45.4 | 257 | US-10-437-963-144036 | Sequence 144036, |
| 34 | 39 | 44.8 | 68 | US-10-437-963-194695 | Sequence 194695, |
| 35 | 39 | 44.8 | 72 | US-10-424-599-165268 | Sequence 165268, |
| 36 | 39 | 44.8 | 91 | US-10-424-599-175273 | Sequence 175273, |
| 37 | 39 | 44.8 | 91 | US-10-424-599-175274 | Sequence 175274, |
| 38 | 39 | 44.8 | 112 | US-10-424-599-249688 | Sequence 249688, |
| 39 | 39 | 44.8 | 117 | US-10-424-599-228673 | Sequence 228673, |
| 40 | 39 | 44.8 | 166 | US-10-424-599-254263 | Sequence 254263, |
| 41 | 39 | 44.8 | 198 | US-10-204-887-96 | Sequence 96, Appl |
| 42 | 39 | 44.8 | 224 | US-10-282-122A-48936 | Sequence 48936, A |
| 43 | 39 | 44.8 | 229 | US-10-424-599-205777 | Sequence 205777, |
| 44 | 39 | 44.8 | 241 | US-10-424-599-255332 | Sequence 255332, |
| 45 | 39 | 44.8 | 313 | US-10-369-493-11799 | Sequence 11799, A |

ALIGNMENTS

RESULT 1

US-10-437-963-169736
; Sequence 169736, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 169736
; LENGTH: 554
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURES:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_68127C.1.pep
; US-10-437-963-169736

Query Match 56.3%; Score 49; DB 16; Length 554;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 YQHDLRAYGFWRL 14
Db 276 FRHDIGWGYWR 287

```

; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9242
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9242

Query Match          49.4%; Score 43; DB 14; Length 256;
Best Local Similarity 72.7%; Pred. No. 63;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 SYQHDRLRAYGF 12
      :|||||:|
Db      224 TYQHDLSAPGF 234
      :|||||:|

RESULT 5
US-09-957-005-11
; Sequence 11, Application US/09957005
; Publication No. US20030104388A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Zhenyu
; APPLICANT: Zhou, Jing
; APPLICANT: Xu, Shuang-yong
; TITLE OF INVENTION: Method For Cloning And Expression of BsmAI Restriction
; TITLE OF INVENTION: Endonuclease And BsmAI Methylase In E. Coli
; FILE REFERENCE: NEB-192
; CURRENT APPLICATION NUMBER: US/09/957,005
; CURRENT FILING DATE: 2001-09-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus A664
US-09-957-005-11

Query Match          49.4%; Score 43; DB 10; Length 465;
Best Local Similarity 53.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      1 RSYQHDRLRAYGF 13
      :|||||:|
Db      214 KSYTKDRRAYEYW 226
      :|||||:|

RESULT 6
US-10-282-122A-59950
; Sequence 59950, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9242
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9242

Query Match          49.4%; Score 43; DB 14; Length 256;
Best Local Similarity 72.7%; Pred. No. 63;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 SYQHDRLRAYGF 12
      :|||||:|
Db      224 TYQHDLSAPGF 234
      :|||||:|

RESULT 5
US-09-957-005-11
; Sequence 11, Application US/09957005
; Publication No. US20030104388A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Zhenyu
; APPLICANT: Zhou, Jing
; APPLICANT: Xu, Shuang-yong
; TITLE OF INVENTION: Method For Cloning And Expression of BsmAI Restriction
; TITLE OF INVENTION: Endonuclease And BsmAI Methylase In E. Coli
; FILE REFERENCE: NEB-192
; CURRENT APPLICATION NUMBER: US/09/957,005
; CURRENT FILING DATE: 2001-09-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus A664
US-09-957-005-11

Query Match          49.4%; Score 43; DB 10; Length 465;
Best Local Similarity 53.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      1 RSYQHDRLRAYGF 13
      :|||||:|
Db      214 KSYTKDRRAYEYW 226
      :|||||:|

RESULT 6
US-10-282-122A-59950
; Sequence 59950, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9242
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9242

Query Match          49.4%; Score 43; DB 14; Length 256;
Best Local Similarity 72.7%; Pred. No. 63;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 SYQHDRLRAYGF 12
      :|||||:|
Db      224 TYQHDLSAPGF 234
      :|||||:|

RESULT 5
US-09-957-005-11
; Sequence 11, Application US/09957005
; Publication No. US20030104388A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Zhenyu
; APPLICANT: Zhou, Jing
; APPLICANT: Xu, Shuang-yong
; TITLE OF INVENTION: Method For Cloning And Expression of BsmAI Restriction
; TITLE OF INVENTION: Endonuclease And BsmAI Methylase In E. Coli
; FILE REFERENCE: NEB-192
; CURRENT APPLICATION NUMBER: US/09/957,005
; CURRENT FILING DATE: 2001-09-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus A664
US-09-957-005-11

Query Match          49.4%; Score 43; DB 10; Length 465;
Best Local Similarity 53.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      1 RSYQHDRLRAYGF 13
      :|||||:|
Db      214 KSYTKDRRAYEYW 226
      :|||||:|

RESULT 6
US-10-282-122A-59950
; Sequence 59950, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9242
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9242

Query Match          49.4%; Score 43; DB 14; Length 256;
Best Local Similarity 72.7%; Pred. No. 63;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 SYQHDRLRAYGF 12
      :|||||:|
Db      224 TYQHDLSAPGF 234
      :|||||:|

RESULT 5
US-09-957-005-11
; Sequence 11, Application US/09957005
; Publication No. US20030104388A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Zhenyu
; APPLICANT: Zhou, Jing
; APPLICANT: Xu, Shuang-yong
; TITLE OF INVENTION: Method For Cloning And Expression of BsmAI Restriction
; TITLE OF INVENTION: Endonuclease And BsmAI Methylase In E. Coli
; FILE REFERENCE: NEB-192
; CURRENT APPLICATION NUMBER: US/09/957,005
; CURRENT FILING DATE: 2001-09-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus A664
US-09-957-005-11

Query Match          49.4%; Score 43; DB 10; Length 465;
Best Local Similarity 53.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      1 RSYQHDRLRAYGF 13
      :|||||:|
Db      214 KSYTKDRRAYEYW 226
      :|||||:|

RESULT 6
US-10-282-122A-59950
; Sequence 59950, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9242
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9242

Query Match          49.4%; Score 43; DB 14; Length 256;
Best Local Similarity 72.7%; Pred. No. 63;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 SYQHDRLRAYGF 12
      :|||||:|
Db      224 TYQHDLSAPGF 234
      :|||||:|

RESULT 5
US-09-957-005-11
; Sequence 11, Application US/09957005
; Publication No. US20030104388A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Zhenyu
; APPLICANT: Zhou, Jing
; APPLICANT: Xu, Shuang-yong
;
```

```
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59950
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-10-282-122A-59950

Query Match          49.4%; Score 43; DB 15; Length 466;
Best Local Similarity 87.5%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      8 RYGFWR 15
Db      430 RYGFWR 437

RESULT 7
US-09-966-997-13
; Sequence 13, Application US/09966997
; Publication No. US20030100052A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Shuang-yong
; APPLICANT: Dore, Andrew
; APPLICANT: Hume, Adam
; APPLICANT: Pelletier, John
; APPLICANT: Zhou, Jing
; TITLE OF INVENTION: Method For Cloning And Expression of Bambi Restriction
; TITLE OF INVENTION: Endonuclease And Bambi Methylase In E. coli And
; TITLE OF INVENTION: Purification Of Bambi Endonuclease
; FILE REFERENCE: NEB-191
; CURRENT APPLICATION NUMBER: US/09/966,997
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus B61
US-09-966-997-13

Query Match          49.4%; Score 43; DB 10; Length 530;
Best Local Similarity 53.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      1 RSYQHDLRAYGF 13
Db      226 RLYNHDRRAFMW 238

RESULT 8
US-10-425-114-59634
; Sequence 59634, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
```

```
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59634
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3606-053-G5_FLI.pep
US-10-425-114-59634

Query Match          48.9%; Score 42.5; DB 15; Length 413;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 10; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY      2 SYQHDRLAY---GFWR 14
Db      107 SYQKSLRAYRACGDWR 122

RESULT 9
US-10-437-963-194505
; Sequence 194505, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 194505
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_90542C.1.psp
US-10-437-963-194505

Query Match          48.3%; Score 42; DB 16; Length 370;
Best Local Similarity 63.6%; Pred. No. 1.3e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 RSYQHDRLAYG 11
Db      182 RYQHDLDLADFG 192

RESULT 10
US-10-767-701-42825
; Sequence 42825, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
```

```
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 42825
; LENGTH: 558
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C2047_1.pep
US-10-767-701-42825

Query Match      48.3%; Score 42; DB 16; Length 558;
Best Local Similarity 41.7%; Pred. No. 1.9e+02;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy      3 YQHDLRAYGFWR 14
Db      167 FRHDINLYMFWK 178

RESULT 11
US-10-437-963-121169
; Sequence 121169, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 121169
; LENGTH: 891
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_24220C.1.pep
US-10-437-963-121169

Query Match      48.3%; Score 42; DB 16; Length 891;
Best Local Similarity 56.2%; Pred. No. 3.1e+02;
Matches 9; Conservative 1; Mismatches 0; Indels 6; Gaps 1;

Qy      6 DLRAYGF-----WRL 15
Db      169 DLRSGYFGILRLWRL 184

RESULT 12
US-10-424-599-211706
; Sequence 211706, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 211706
```

```
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_WRT3847_33198C.1.pep
US-10-424-599-211706

Query Match      47.1%; Score 41; DB 15; Length 129;
Best Local Similarity 42.9%; Pred. No. 69;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy      2 SYQHDLRAYGFWR 15
Db      49 NWDPNLRALGFWKI 62

RESULT 13
US-10-293-958-5
; Sequence 5, Application US/10293958
; Publication No. US20030140380A1
; GENERAL INFORMATION:
; APPLICANT: HENKES, STEFAN
; APPLICANT: CHEN, RUOYING
; APPLICANT: DA COSTA E SILVA, OSWALDO
; APPLICANT: VAN THIELEN, NOCHA
; TITLE OF INVENTION: AMINE OXIDASE STRESS-RELATED POLYPEPTIDES AND METHODS
; TITLE OF INVENTION: OF USE IN PLANTS
; FILE REFERENCE: 16313-0174
; CURRENT APPLICATION NUMBER: US/10/293,958
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 60/345,969
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Brassica napus
US-10-293-958-5

Query Match      47.1%; Score 41; DB 14; Length 497;
Best Local Similarity 54.5%; Pred. No. 2.5e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy      3 YQHDLRAYGFWR 13
Db      115 YDHLESYGLY 125

RESULT 14
US-10-106-275-6
; Sequence 6, Application US/10106275
; Publication No. US20030186363A1
; GENERAL INFORMATION:
; APPLICANT: Zhu, Zhenyu
; APPLICANT: Xu, Shuang-yong
; TITLE OF INVENTION: Method For Cloning And Expression of BsaI Restriction Endonuclease
; TITLE OF INVENTION: BsaI Methylase In E. Coli
; FILE REFERENCE: NEB-202
; CURRENT APPLICATION NUMBER: US/10/106,275
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus 6-55
US-10-106-275-6

Query Match      47.1%; Score 41; DB 14; Length 544;
Best Local Similarity 53.8%; Pred. No. 2.8e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

QY 1 RSYQHDLRAYGFW 13
|||:|:|:
Db 229 RSYTDRRAFYW 241

RESULT 15
US-10-437-963-129380
; Sequence 129380, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 129380
; LENGTH: 573
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_31641C.1.pap
US-10-437-963-129380

Query Match 47.1%; Score 41; DB 16; Length 573;
Best Local Similarity 50.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 RSYQHDLRAYGFW 14
|||:|:|:
Db 263 RSYRHGLSSGGWK 276

RESULT 16
US-10-389-566-1468
; Sequence 1468, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)D
; CURRENT APPLICATION NUMBER: US/10/389,566
; CURRENT FILING DATE: 2003-03-31
; PRIOR FILING DATE: 2002-03-15
; PRIOR FILING DATE: 2002-03-15
; PRIOR FILING DATE: 2002-06-25
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2459
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1468
; LENGTH: 692
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-10-389-566-1468

Query Match 47.1%; Score 41; DB 15; Length 692;
Best Local Similarity 57.1%; Pred. No. 3.5e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 RSYQHDLRAYGFW 14
|||:|:|:
Db 236 RSFQFLRTYPEWR 249

RESULT 17
US-10-389-566-1481
; Sequence 1481, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)D
; CURRENT APPLICATION NUMBER: US/10/389,566
; CURRENT FILING DATE: 2003-03-31
; PRIOR FILING DATE: 2002-03-15
; PRIOR FILING DATE: 2002-03-15
; PRIOR FILING DATE: 2002-06-25
; PRIOR FILING DATE: 2002-06-26
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2459
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1481
; LENGTH: 817
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-10-389-566-1481

Query Match 47.1%; Score 41; DB 15; Length 817;
Best Local Similarity 57.1%; Pred. No. 4.1e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 RSYQHDLRAYGFW 14
|||:|:|:
Db 362 RSFQFLRTYPEWR 375

RESULT 18
US-10-437-963-181030
; Sequence 181030, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 181030
; LENGTH: 1479
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_78343C.1.pap
US-10-437-963-181030

Query Match 47.1%; Score 41; DB 16; Length 1479;
Best Local Similarity 50.0%; Pred. No. 7.3e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 RSYQHDLRAYGFW 14
|||:|:|:
Db 263 RSYRHGLSSGGWK 276

RESULT 19

US-10-424-599-281147
; Sequence 281147, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 281147
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_95899C.1.pgp
US-10-424-599-281147

Query Match 46.0%; Score 40; DB 15; Length 76;
Best Local Similarity 63.6%; Pred. No. 60;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 5 HDLRAYGFWR 15
||:|||||
Db 1 HDDCRWGFWR 11

RESULT 20
US-10-156-761-12936
; Sequence 12936, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12936
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12936

Query Match 46.0%; Score 40; DB 14; Length 150;
Best Local Similarity 63.6%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RSYQHDLRAYG 11
||:|||||
Db 86 RAYESSLRAYG 96

RESULT 21
US-09-765-815-13
; Sequence 13, Application US/09765815
; Publication No. US20030166623A1
; GENERAL INFORMATION:
; APPLICANT: Balk, Steven

; TITLE OF INVENTION: NO. US20030166623A1el Steroid Hormone Receptor
; TITLE OF INVENTION: Interacting Protein Kinase
; FILE REFERENCE: 01948/068002
; CURRENT APPLICATION NUMBER: US/09/765,815
; CURRENT FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/176,859
; PRIOR FILING DATE: 2000-01-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-765-815-13

Query Match 46.0%; Score 40; DB 10; Length 151;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RSYQHDLRAYGF 12
||:|||||
Db 52 RYQHEVEMYNF 63

RESULT 22
US-10-380-132-18
; Sequence 18, Application US/10380132
; Publication No. US20030182679A1
; GENERAL INFORMATION:
; APPLICANT: SunGene GmbH & Co. KGaA
; TITLE OF INVENTION: Improved processes for vitamin E biosynthesis
; FILE REFERENCE: NAE445/2000
; CURRENT APPLICATION NUMBER: US/10/380,132
; CURRENT FILING DATE: 2003-03-11
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-380-132-18

Query Match 46.0%; Score 40; DB 14; Length 223;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YQHDLRAYGWR 14
||:|||||
Db 8 YQAKLKLYSYWR 19

RESULT 23
US-09-287-599-2
; Sequence 2, Application US/09287599
; Patent No. US20020151071A1
; GENERAL INFORMATION:
; APPLICANT: Handelsman, Jo
; APPLICANT: Klimowicz, Amy
; TITLE OF INVENTION: Enterotoxin-Deficient Bacillus
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/287,599
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 960296
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 334 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-287-599-2

Query Match 46.0%; Score 40; DB 9; Length 334;
Best Local Similarity 63.6%; Pred. No. 2.5e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SYQHDRLRAYGF 12
DB 181 SIGHDVRAEFP 191

RESULT 24
US-10-425-114-47984
; Sequence 47984, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 47984
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700787750_FLI.pep
US-10-425-114-47984

Query Match 46.0%; Score 40; DB 15; Length 478;
Best Local Similarity 54.5%; Pred. No. 3.5e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 YQHDRLRAYGF 13
DB 102 YDHDLESYGLF 112

RESULT 25
US-10-424-599-166245
; Sequence 166245, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 166245
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_121134C.1.pep
US-10-424-599-166245

Query Match 46.0%; Score 40; DB 15; Length 490;
Best Local Similarity 54.5%; Pred. No. 3.6e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 YQHDRLRAYGF 13
DB 114 YDHDLESYGLF 124

RESULT 26
US-10-295-403-166
; Sequence 166, Application US/10295403
; Publication No. US20030101481A1
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc
; APPLICANT: Broun, Pierre
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, Lynne
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddle, James
; APPLICANT: Zhang, James
; APPLICANT: Benito, Maria-Ines
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Fromm, Mike
; TITLE OF INVENTION: PLANT GENE SEQUENCES I
; FILE REFERENCE: MBI-0003
; CURRENT APPLICATION NUMBER: US/10/295,403
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US/09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 60/101,349
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: 60/103,312
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: 60/108,734
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/113,409
; PRIOR FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 166
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G849
US-10-295-403-166

Query Match 46.0%; Score 40; DB 14; Length 618;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRAYGF 13
DB 583 LRAYGW 589

RESULT 27
US-09-934-455-142
; Sequence 142, Application US/09934455

Publication No. US20030121070A1
; GENERAL INFORMATION:
; APPLICANT: Adam, Luc
; APPLICANT: Creelman, Robert
; APPLICANT: Dubell, Arnold
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddle, James
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Reuber, Lynne
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Pineda, Omaira
; TITLE OF INVENTION: Genes for Modifying Plant Traits IV
; FILE REFERENCE: MBI-0025
; CURRENT APPLICATION NUMBER: US/09/934,455
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227439
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: MBI-0022
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: MBI-0023
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 516
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 142
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-934-455-142

Query Match 46.0%; Score 40; DB 10; Length 619;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRAYGFW 13
Db 583 LRAYGYW 589

RESULT 28
US-10-225-066A-894
; Sequence 894, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBELL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: PILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROWN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI0036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/837,444
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14

; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 894
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-225-066A-894

Query Match 46.0%; Score 40; DB 15; Length 619;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRAYGFW 13
Db 583 LRAYGYW 589

RESULT 29
US-10-374-780A-164
; Sequence 164, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddle, James
; APPLICANT: Brown, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 164
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G849 (conserved domain in AA coordinates: 324-413, 504-583)
US-10-374-780A-164

Query Match 46.0%; Score 40; DB 15; Length 619;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRAYGFW 13
|||||:
Db 583 LRAYGYW 589

RESULT 30

US-10-412-699B-566
; Sequence 566, Application US/10412699B
; Publication No. US20040045049A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Zhang, James
; APPLICANT: Fromm, Michael E.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Kiechmann, Jose Luis
; APPLICANT: Adam, Luc J.
; APPLICANT: Broun, Pierre E.
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddle, James S.
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Creslan, Robert A.
; APPLICANT: DuBell, Arnold N.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Kumimoto, Roderick
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI-0048CIP
; CURRENT APPLICATION NUMBER: US/10/412,699B
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 566
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G849
US-10-412-699B-566

Query Match 46.0%; Score 40; DB 15; Length 619;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRAYGFW 13
|||||:
Db 583 LRAYGYW 589

RESULT 31

US-10-437-963-183525
; Sequence 183525, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 183525
; LENGTH: 677
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_80608C.1.pep
US-10-437-963-183525

Query Match 46.0%; Score 40; DB 16; Length 677;
Best Local Similarity 61.5%; Pred. No. 5e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SYQHDLRAYGFW 14
|||||:
Db 525 SYLNDLVNYGFW 537

RESULT 32

US-09-934-868-52
; Sequence 52, Application US/09934868
; Patent No. US20020137190A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odom, James M.
; APPLICANT: Schenzle, Andreas J
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1596 US NA
; CURRENT APPLICATION NUMBER: US/09/934,868
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 52
; LENGTH: 920
; TYPE: PRT
; ORGANISM: Methylobionas 16a
; FEATURE:
; OTHER INFORMATION: Nasa
US-09-934-868-52

Query Match 46.0%; Score 40; DB 9; Length 920;
Best Local Similarity 50.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YOHDLRAYGFW 14
|||||:
Db 727 YHVIRGHGFW 738

RESULT 33

US-10-437-963-144036
; Sequence 144036, Application US/10437963
; Publication No. US20040123343A1

QY 3 YOHDLRAYGFW 14
|||||:
Db 727 YHVIRGHGFW 738

; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 144036
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_44888C.1.pap
US-10-437-963-144036

Query Match 45.4%; Score 39.5; DB 16; Length 257;
Best Local Similarity 56.2%; Pred. No. 2.3e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 2 SYQHDRLAY---GFWR 14
|||:|:|:|:|
Db 222 SYQSLKAYRDCGDWR 237

RESULT 34
US-10-437-963-194695
; Sequence 194695, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 194695
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_90715C.1.pap
US-10-437-963-194695

Query Match 44.8%; Score 39; DB 16; Length 68;
Best Local Similarity 42.9%; Pred. No. 78;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 SYQHDRLAYGFWRLL 15
|:|:|:|:|
Db 9 SWGHTIRGYAIWAL 22

RESULT 35
US-10-424-599-165268
; Sequence 165268, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 165268
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_120252C.1.pap
US-10-424-599-165268

Query Match 44.8%; Score 39; DB 15; Length 72;
Best Local Similarity 46.7%; Pred. No. 82;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 RSYQHDLRAYGFWRLL 15
|||:|:|:|:|
Db 16 RSYGNILDPFGIWHL 30

RESULT 36
US-10-424-599-175273
; Sequence 175273, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 175273
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_129290C.1.pap
US-10-424-599-175273

Query Match 44.8%; Score 39; DB 15; Length 91;
Best Local Similarity 45.5%; Pred. No. 1e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 YQHDLRAYGFW 13
:|:|:|:|
Db 44 WSYPMRPYGEW 54

RESULT 37
US-10-424-599-175274
; Sequence 175274, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 175274
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_129291C.1.pep
US-10-424-599-175274

Query Match 44.8%; Score 39; DB 15; Length 91;
Best Local Similarity 45.5%; Pred. No. 1e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 YQHDLRAYGF 13
Db 44 WSPYMRPYGF 54

RESULT 38
US-10-424-599-249688
; Sequence 249688, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 249688
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_67498C.1.pep
US-10-424-599-249688

Query Match 44.8%; Score 39; DB 15; Length 112;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 HDLRAYGF 13
Db 4 HDLATYGAW 12

RESULT 39
US-10-424-599-228673
; Sequence 228673, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 228673
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(117)
; OTHER INFORMATION: unsure at all Xaa locations

; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_48519C.1.pep
US-10-424-599-228673

Query Match 44.8%; Score 39; DB 15; Length 117;
Best Local Similarity 54.5%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 5 HDLRAYGF 15
Db 34 HDIRSYVLWLL 44

RESULT 40
US-10-424-599-254263
; Sequence 254263, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 254263
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(166)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_71623C.1.pep
US-10-424-599-254263

Query Match 44.8%; Score 39; DB 15; Length 166;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 HDLRAYGF 13
Db 110 HDLFNFGFW 118

Search completed: March 18, 2005, 18:43:34
Job time : 137 secs

This Page Blank (uspto)